الدكتور علي عبد جاسم الزاملي
الدكتور عبدالله بن محمد الصارمي
الدكتور علي مهدي كاظم

مفاهيم وتطبيقات في التقويم والقياس التربوي



مفاهيم وتطبيقات

في التقويم والقياس التربوي


بسم الله الرحمن الرحيم

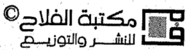
الدكتور علي عبد جاسم الزاملي
جامعة السلطان قابوس

الدكتور عبدالله بن محمد الصارمي
وكيل وزارة التعليم العالي

الدكتور علي مهدي كاظم
جامعة السلطان قابوس

مفاهيم وتطبيقات

في التقويم والقياس التربوي



الطبعة الأولى 1429 هـ 2009 م

مكتبة الفلاح©
للنشر والتوزيع

دولة الكويت
حولي-شارع بيروت - عمارة الأطباء
هاتف: 1985 264 فاكس: 7784 264 965+
ص.ب: 4848 الصفاة-الرمز البريدي 13049

دولة الإمارات العربية المتحدة
العين: - ص.ب 16431 هاتف: 7662189 فاكس: 7657901 3 971+
دبي: - ص.ب: 20438 هاتف: 2630618 فاكس: 2630628 4 971+

جمهورية مصر العربية
37 شارع النصر - امتداد رمسيس2
مقابل وزارة المالية ومصلحة الجمارك
مدينة نصر - القاهرة
تلفاكس 8143 262 202+
E-mail: alfalah_egypt@ hotmail.com

المملكة الأردنية الهاشمية
دار حنين للنشر والتوزيع
العبدلي- مقابل البنك العربي- عمارة الددو
هاتف 5695611 فاكس 1208 568 6 962+
ص.ب 927385 الرمز البريدي 11190
عمان - الأردن
e-mail: dar_honin@yahoo.com

الإهداء

إلى كل من يُفني نفسهُ ليبدد الظلام...

إلى كل معلمٍ ومعلمة...... وكل باحثٍ وباحثه.......

نهدي هذا الكتاب.

المحتويات

الباب الأول

التقويم

الفصل الأول: التطور التاريخي لحركة التقويم

الفصل الثاني: مفاهيم أساسية

الفصل الثالث: تقويم البرامج التربوية باستخدام النماذج

الفصل الرابع: تخطيط التقويم ومعاييره

الفصل الخامس: أدوات التقويم

الباب الثاني

القياس

الفصل السادس: مفهوم القياس

الفصل السابع: اختبارات التحصيل وأنواعها

الفصل الثامن: التقويم البديل لاختبارات الورقة والقلم

الباب الثالث

الإحصاء والبرامج الحاسوبية

الفصل التاسع: استخدام الإحصاء في القياس والتقويم

الفصل العاشر: البرمجيات المستخدمة في القياس والتقويم

المقدمة

حظي موضوع القياس والتقويم التربوي والنفسي باهتمام كثير من علماء التربية وعلم النفس والمؤسسات التربوية الخاصة بإعداد المعلمين الذين ساهموا مساهمة فاعلة في تطوير نظريات القياس والتقويم وتصميم الاختبارات النفسية والتربوية والاجتماعية، وإهتمت بعض المؤسسات التربوية بفتح برامج دراسات عليا يتخصص فيها الطالب في هذا المجال. كما قامت تلك المؤسسات أيضا بتنفيذ العديد من الورش التدريبية لسد حاجة المعلمين الماسة لأساسيات هذا المجال لا سيما عملية تصميم أو بناء اختبارات التحصيل المقننة التي يجهلها بعض المعلمين، والاحتفاظ بنماذج من الأسئلة ذات المستوى الجيد من حيث معامل الصعوبة ومعامل التمييز، علما بأن الكثير من مدارس الدول المتقدمة وبعض مدارس الدول العربية أصبحت تهتم بإيجاد ما يسمى ببنك الأسئلة.

ولم يقتصر اهتمام الباحثين والمؤسسات التربوية على مجال القياس فقط، بل أن التقويم أيضا أخذ نصيباً كبيرا من ذلك الاهتمام، ولم يعد التقويم في أضيق حالاته عندما كان يقتصر على تقويم تحصيل الطالب، بل امتد الى تقويم الخطط التعليمية وتقويم المدارس (ذاتيا وخارجيا) وتقويم البرامج والمؤسسات التربوية الأخرى من خلال جمع بيانات عن أدائها وتحديد نقاط القوة والضعف فيها ومن ثم وضع خطط التطوير.

وانطلاقا مما تقدم يضع المؤلفون هذا الكتاب بين يدي القارئ العربي لا سيما المتخصص ويحدوهم الأمل بأنه سيلبي حاجاته في هذا المجال، وقد كُتبت موضوعاته بأسلوب واضح ومبسط مدعوما بأمثلة تطبيقية مراعيا جميع المستويات من الباحثين والمعلمين إلى طلبة الدراسات الأولية والعليا في الجامعات.

ولأجل إظهار الكتاب بالمظهر المناسب لحاجات هؤلاء، شمل الكتاب المفاهيم النظرية والأساليب التطبيقية للقياس والتقويم واهتدى بالبرامج الحاسوبية التي سهّلت عمليات القياس والتقويم وأفرد لها فصلا خاصا يساعد الطلبة على تحليل البيانات ونتائج الاختبارات وتفسيرها، وقد ضم الكتاب ثلاثة أبواب، هي :

الباب الأول: التقويم، ويتألف من خمسة فصول:

- الفصل الأول: التطور التأريخي لحركة التقويم.
- الفصل الثاني: مفاهيم أساسية .
- الفصل الثالث: تقويم البرامج التربوية.
- الفصل الرابع: تخطيط التقويم ومعاييره.
- الفصل الخامس: أدوات التقويم.

الباب الثاني: القياس، ويتألف من ثلاثة فصول هي:

- الفصل السادس: مفهوم القياس.
- الفصل السابع: اختبارات التحصيل وأنواعها.
- الفصل الثامن: اختبارات التقويم البديل.

الباب الثالث: الإحصاء والبرمجيات الحاسوبية ويتألف من فصلين هما:

- الفصل التاسع: استخدام الإحصاء في القياس والتقويم.
- الفصل العاشر: البرامج الحاسوبية المستخدمة في القياس والتقويم.

وأخيرا، نأمل أن ينال جهدنا المتواضع هذا رضا التربويين على اختلاف مستوياتهم ومن الله التوفيق.

المؤلفون

مسقط 2008

الباب الأول

التقويم Evaluation

مفاهيم وتطبيقات

الفصل الأول

التطور التاريخي لحركة التقويم

بدايات التقويم في الصين واليابان

التقويم في أوربا

التقويم في الولايات المتحدة

التقويم في الوطن العربي

تجربة سلطنة عُمان

1

الفصل الأول

التطور التاريخي لحركة التقويم

قبل الخوض في تحديد مفهوم التقويم وأهميته في العملية التربوية لابد لنا من إلقاء نظرة موجزة على الجذور التاريخية لتطور حركة التقويم والاختبار في العالم. إذ مما لا شك فيه أن أية مقدمة تذكر عن مجال معين تستوجب إعطاء بعض المعالم البسيطة عن تطور ذلك المجال بإلقاء الضوء على المراحل التي مر بها والتطورات المختلفة والنظريات والأفكار التي أدت بتراكمها التاريخي إلى استكمال صورته بوضعها الحاضر.

بدايات التقويم في الصين واليونان

إن تاريخ التقويم في الحقيقة أطول بكثير مما يتصور البعض عندما يعيدون جذوره إلى مرحلة زمنية قريبة. إن مفهوم تقويم الأفراد والبرامج كان معروفاً منذ حوالي (2000 سنة قبل الميلاد) عندما كان الصينيون يستخدمون في نظامهم التعليمي امتحانات خاصة لاختيار الأفراد لبعض الخدمات والوظائف المدنية العليا. وفي عصر النهضة اليونانية (500 ق.م) كان المعلمون ومنهم (سقراط) يستخدمون أساليب تقويمية شفوية كجزء مهم من عملية التعلم والتعليم. ويبدو أن أساليب التقويم الشفوية التي استخدمت في الماضي البعيد امتدت صلتها بأساليب التقويم التي سادت القرنين الماضيين، عندما شهدت المدارس الرسمية التقليدية استخدام ما يسمى بطريقة (التسميع الشفهي، Oral Citation) التي كان الهدف منها تدريب التلاميذ على حفظ حقائق أو مقطوعات معينة وإعادتها من الذاكرة. لذلك أصبح من الطبيعي أن يلجأ المعلمون إلى التسميع الشفهي للتثبت من مدى تحقق هذا الهدف وبذلك أصبحت هذه الطريقة من الأساليب السائدة في تقويم الطلبة. ويُعزي البعض شيوع هذا الأسلوب في التقويم إلى أن أهداف التعلم كانت بسيطة ومحددة من جهة، وغياب المواد الكتابية التي لم

تكن متيسرة دوماً كما هي اليوم من جهة أخرى. إضافة إلى أن استخدام المواد الكتابية كان صعباً ولم تشهد تبسيطاً لطرق استخدامها حتى عقودٍ قريبة العهد نسبياً.

لقد تطور مفهوم التسميع الشفهي بعد ذلك ليعكس ظله على أسلوب مشابه لكنه مطور نوعاً ما، وهو أسلوب (الامتحانات الشفهية Oral Examinations) التي كانت حتى عهد قريب يعتمد عليها في تقويم الطلبة والتي تقوم على السؤال والجواب أو علًى نوع من الحوار بين الممتحن (المعلم) والطالب وكما نشهده اليوم في مناقشات رسائل الماجستير والدكتوراه لنيل لقبٍ علمي جامعي معين وهو تقليد نشأ عن التعليم في القرون الوسطى. ولا يزال الامتحان الشفهي العنصر المميز لأساليب التقويم التربوي في بعض بلدان العالم.

التقويم في أوروبا

في أنحاء من أوربا أنشئ في أوائل القرن التاسع عشر ما يشبه نظام الامتحان المركزي بهدف تحقيق المزيد من المساواة والعدالة في التقويم. في فرنسا بقي هذا الامتحان شفهياً لمدة طويلة، ولم يتحول في انكلترا إلى امتحان كتابي حتى أواسط القرن التاسع عشر. وفي هذه الفترة أخذ المربون يدركون مساوئ الاعتماد الكلي على الامتحانات الشفهية وما يسمى بأسلوب التسميع الشفهي عندما بدأ الاهتمام باستخدام الاختبارات التحريرية (الكتابية) لأجل دعم نتائج الطرق المعتمدة على التسميع الشفهي، وزاد الإقبال من قبل المعلمين على استخدام الاختبار الكتابي بعد ازدياد وسهولة الحصول على المواد الكتابية المناسبة، وكانت هذه الاختبارات من النوع الذي يتكون من مجموعة من المسائل الرياضية أو الأسئلة التي تتطلب إجابات من نوع المقال.

إن ظهور الامتحانات التحريرية لم يؤد إلى إلغاء أسلوب الامتحانات الشفهية من الأنظمة التعليمية، إذ أنها لا زالت تمثل أحد الأساليب التقويمية المستخدمة في تقويم سلوك الطالب من جوانب عديدة ولو أنها (الامتحانات

الشفهية) أصبحت أقل انتشاراً أو قلّ الاعتماد عليها بسبب ظهور الاختبار الكتابي الذي أضاف إلى أسلوب التقويم القديم أسلوباً آخر.

وفي نهاية القرن التاسع عشر وبداية القرن العشرين شهد ميدان التقويم والاختبار تطوراً جديداً عندما ظهرت إلى الوجود (اختبارات الذكاء Intelli-gence Tests) على أثر الأبحاث التي قام بها بعض علماء النفس والتربية ومنهم العالم الفرنسي (الفرد بينية Elferd Binet) عام 1905 والذي قيل عن اختباره أنه لا يوجد اختبار في العالم نال عناية علماء النفس كالعناية التي نالها هذا الاختبار الذي يتألف من (30 مشكلة) أو سؤالاً تتدرج في مستوى صعوبتها. لقد صممت أسئلة الاختبار لتغطي أنواعاً متعددة من الوظائف العقلية مع تأكيد خاص على الحكم Judgment والاستيعاب Comprehension والاستدلال Reasoning، والتي اعتبرها (بينية) المكونات الأساسية للذكاء وقد طبق هذا المقياس في مرحلة تصميمه على بعض التلاميذ المتأخرين عقلياً وبعض الراشدين في المدارس الفرنسية.

التقويم في الولايات المتحدة الأمريكية

في عام 1916 نُشرت في الولايات المتحدة صورةً معدلة لاختبار الذكاء هذا عرفت باسم (اختبار ستانفورد - بينية) نسبة إلى الجامعة التي كان يعمل فيها العالم الشهير (تيرمان، Terman) وهو الذي طوّر هذا المقياس الفرنسي الأصل.

لقد كان اختبار ستانفورد - بينية المفتاح الذي فتح الباب للاختبارات الموضوعية لتظهر منه على العالم والتي لازالت تستعمل في كثير من المدارس في الوقت الحاضر. ولا يُنكر أن هذه الاختبارات ساهمت في تطوير طرائق التدريس التي يستخدمها المعلمون وأساليب التقويم الخاصة بطلبتهم عندما أضافت طرقاً أخرى في هذين المجالين. ولابد لنا هنا من أن نذكر إن اتساع حركة اختبارات الذكاء والاستعدادات في مطلع القرن العشرين رافقها أو تبعها نشوء حركة بناء (الاختبارات التحصيل المقننة، Standardized Achievement Tests) واتساع مفهوم القياس والتقويم في التربية والتعليم بفضل جهود عدد

من العلماء البارزين من أمثال "مان Man ورايز Rice وثورندايك Thorndike الذي يعتبر بحق الرائد الحقيقي للقياس والتقويم ويعود له الفضل في وضع أسس بناء الاختبارات الموضوعية وما نتج عن ذلك من انتشار هذه الاختبارات لقياس التغير الإنساني وهنا ازدهرت حركة استخدام القياس لتحديد القابليات الإنسانية واقترن مفهوم التقويم بمفهوم القياس في التربية وعلم النفس.

إن تطور حركة القياس هذه شجعت المعلمين على بناء وتطوير اختبارات التحصيل التي أصبحت تشكل أساساً مهماً في نظام التقويم والقياس في المدارس.

أما أساليب قياس الشخصية والرغبات فإنها نمت وتعاظمت خلال هذه الفترة أيضاً حيث أخذت المؤسسات الصناعية والعسكرية باستخدام هذه الأساليب لتقويم المرشحين للانخراط فيها أو أنها استخدمت كجزء مهم من أساليب الاختيار والتصنيف للأفراد الراغبين في الانخراط في تلك المؤسسات.

ومن الجدير بالذكر أن حركة الاختبارات التحصيلية المقننة أثارت الحماس والاهتمام لدى المعلمين ودفعتهم إلى المغالاة في التشديد على استخدامها وإساءة تفسيرها حتى أنهم راحوا يستخدمونها دون مراعاة للأهداف التي وضعت من أجلها، وأصبح إجراؤها وكأنه غاية في حد ذاته، بل أن هذه الاختبارات استهوت الكثير منهم ودفعتهم إلى الاعتقاد بأنها المرحلة النهائية في تاريخ تطور أساليب التقويم والتركيز عليها وكأنها الطريقة الوحيدة الصالحة للاستخدام في تقويم تحصيل تلاميذهم.

وخلال العقد الرابع من القرن الماضي شهد ميدان التربية والتعليم تطورات جديدة أثرت في توسيع وتطوير مفهوم التقويم وتطبيقاته، إن كان على مستوى تقويم الطلبة أو تقويم البرامج والمشاريع التربوية، ومن هذه التطورات ما قام به العالم الشهير (تايلر، Tylor) من تطبيق لآرائه التربوية في مجالي المناهج والتقويم. لقد عمل تايلر قرابة الخمسة عشر سنة لتطوير أعماله ووجهة نظره حتى توصل إلى أسلوب كان مغايراً لوجهات النظر المطروحة في تلك

الفترة عندما ركز جهوده وآرائه في الأهداف التربوية. لقد حدد تايلر مفهوم التقويم كعملية تقرير أو تحديد لمدى تحقق أو عدم تحقق الأهداف التربوية التي عرّفها بأنها: تغيرات مطلوب إحداثها لدى الطلبة.

لقد كان لأسلوب تايلر في التقويم تأثيراً كبيراً في تخطيط دراسات تقويمية للسنوات الثلاثين اللاحقة لظهوره ولازالت حتى اليوم تؤثر في عملية التخطيط والتقويم لاسيما في مجال المناهج. كما اشتهرت حركة تايلر في الدراسة التي قام بها هو وسمث Smith ونشرت عام 1942 في الولايات المتحدة الأمريكية، والتي سميت (دراسة الثمان سنوات) حيث صممت لاختبار مدى فعالية المناهج المستخدمة واستراتيجيات التدريس في عدد من مدارس الولايات المتحدة.

لقد ساعدت تلك الدراسة الرائدة تايلر لتوسيع واختبار وإظهار آرائه في التقويم التربوي، كما ساعدت على نمو استخدام وانتشار الاختبارات والمقاييس وقوائم التقدير والاستفتاءات على مدى واسعا في ميدان التربية والتعليم.

لقد شهدت حركة التقويم - على أثر أعمال تايلر وزملاؤه- تطوراً ملحوظاً لاسيما بما أحرز من تقدم نتيجة للجهود العلمية المنظمة لتحديد وتصنيف الأهداف التربوية والتي كان من أبرزها (تصنيف الأهداف المعرفية Cognitive Domain) الذي نشره بلوم عام 1956، و(تصنيف الأهداف الوجدانية Affective Domain) الذي نشره كراثهول عام 1968، بالإضافة إلى الجهود التي بذلها كل من ديف عام Dave 1968 وهارو عام 1972 لتحديد الأهداف (النفسحركية Psychomotor Domain).

لقد صاحب عملية تصنيف الأهداف محاولات جادة لصياغتها في عبارات سلوكية تحدد نتائج التعلم التي يمكن ملاحظتها وقياسها كتلك التي قام بها تايلر Tylor 1956 وميجر Mager, 1962، وجانيية Gayne, 1963 وغيرهم.

ومن المؤشرات الأخرى لنمو حركة التقويم في القرن العشرين بروز حركة الاعتماد أو التصديق في التقويم (Accreditation) التي رسخت أقدامها

وتعاظمت آثارها في الميدان التربوي، وبرزت المؤسسات الرسمية الخاصة بإجراء هذه المهمة لاختبار مدى موثوقية المدارس والمعاهد والكليات والمؤسسات التعليمية الموجودة في الولايات المتحدة من خلال أساليب تقويمية تستخدمها مؤسسات الاعتماد لاختبار فاعلية وصدق وجدارة تلك المؤسسات لتكون أهلاً للمهمة التي تقوم بها، وبذلك تنامي وتطور مفهوم التقويم التربوي ليخرج من مجاله الضيق الذي بدأه عندما اقتصر على تقويم تحصيل الطالب فقط إلى ميادين تربوية أوسع من ذلك. وبالإضافة إلى ذلك ظهرت أيضاً وتنامت بعض المؤسسات الخاصة بتقديم خدمات التقويم والاختبار للمستفيدين أو المحتاجين إليها، ومن ذلك ما يسمى (خدمات الاختبار التربوي Educational Testing Service) التي وجدت عام 1947 و(مؤسسة كارنيجي Carrnegie Corporation) و(مجلس التربية الأمريكي American Council of Education) وغيرها من مؤسسات التقويم التي تقدم خدمات مرموقة في هذا المجال.

وفي فترة الخمسينات وأوائل الستينات من القرن الماضي ظهرت أنواع جديدة ومستحدثات تربوية خاصة بالمناهج، ومعها ظهرت الحاجة لوسائل تقويمية جديدة. وحاول العديد من المقومين استخدام نماذج تقويمية متنوعة. وخلال العقدين السابع والثامن من القرن نفسه توّسع مفهوم التقويم إلى مجالات عديدة مثل الإدارة، الصناعة، الجيش، الاقتصاد والاجتماع وامتد في مجال التربية إلى معالجة أمور كثيرة منها ما يتعلق بالطالب، الكتاب، المدرسي والمشاريع والخطط والبرامج التربوية. واشتهر في هذه الفترة مقومون أحدثوا الكثير من الجدل بما طرحوه في ميدان التقويم وما تركوه من تأثير كبير في مفهومه ونذكر منهم على سبيل المثال: كرونباخ Cronbach، وسكرفن Secrevine، وستيك Steak، ولندفل Lindvall، وكوبا Guba، وستيفلبيم Stufflebeam وغيرهم. كما تزايد ظهور العديد من المجلات التي تهتم بأمور التقويم والتي ساعدت في توضيح وتطوير مفهومه واستخداماته ومنها (مجلة القياس التربوي Journal of Educational Measuremant)، ومجلة (التقويم

التــربوي وتحليل الســيــاســة Educational Evaluation and Policy Analysis ومـجلة (دراسـات في التـقـويم Studies in Evaluation) ومـجلة (التـقـويم وتخطيط البرامج Evaluation and Program Planning) ومجلة (أخبار التقويم Evaluation News) ومجلات أخرى مثل: (Evaluation Review و Evalua-tion Quarterly) وغيرها.

لقد أصبح مفهوم التقويم يؤكد على استخدام أساليب تقويمية متنوعة تركز على واقع العملية التربوية وعناصرها المهمة من جهة، وما يوضع لها من أهداف محددة من جهة أخرى.

التقويم في الوطن العربي

أما حركة التقويم في الوطن العربي فإنها لم تكن بعيدة عما كان حولها في بلدان العالم المتقدم، ولم تشهد نوعاً من التطور الرسمي إلا في حدود الستينات من القرن العشرين على الرغم من ومضات بسيطة حدثت في الثلاثينات عندما توجه الاهتمام بالامتحانات واقتراح إدخال أساليب جديدة في أسس التقويم ومصادره، واعتماد البطاقة المدرسية والملف الشخصي للطالب لاسيما في مصر والعراق.

وكانت أول دراسة جدية للامتحانات في البلاد العربية تمَّت عام 1961 عندما انعقد في بيروت مؤتمر تربوي ضم العديد من ممثلي الدول العربية حيث كان الهدف منه دراسة واقع الامتحانات المدرسية العربية والخروج بتوصيات موحدة بهذا الشأن. تبع ذلك المؤتمر الثقافي العربي السادس الذي عقد في الجزائر عام 1964، حيث تناول (دراسة نظم الامتحانات المدرسية ودورها في تقويم الطلاب وتوجيههم)، وكذلك (حلقة توحيد أنظمة الامتحانات والانتقال في المراحل الدراسية المختلفة) التي عقدت في القاهرة عام 1970 و(اجتماع خبراء تطوير نظم الامتحانات في البلاد العربية) الذي عقد في الكويت عام 1974 . لقد أكدت هذه الاجتماعات على جملة من التوصيات منها ضرورة شمول عملية التقويم جميع جوانب شخصية الطالب وتأكيدها

على مسؤولية معلم الصف في إجراء تقويمات مستمرة لطلبته خلال العام الدراسي يكون لها دور في التقويم النهائي للطالب.

ومن الأمور الأخرى المهمة التي وردت في هذه التوصيات التأكيد على وضع مجموعة من الاختبارات المقننة في المواضيع الدراسية المختلفة واستخدام أنواع من الاختبارات النفسية مثل اختبارات الذكاء، والاستعداد والميول لتقويم جوانب معينة من شخصية الطالب، كما أكدت على ضرورة تدريب المعلمين على إعداد الاختبارات التحصيلية والأساليب الحديثة في تقويم التلاميذ ومتابعتهم.

كما أكدت بعض البلدان العربية على وزارات التربية بضرورة الاهتمام بعملية تقويم الطلبة اهتماماً واضحاً من خلال استخدام المعلمين والمدرسين لأكثر من أسلوب لتقويم شخصية الطالب وتحصيله، ووضع برامج وخطط لتدريب المعلمين والمدرسين بغية إطلاعهم وتدريبهم على صياغة الأسئلة الامتحانية والتنوع في أساليب التقويم. وفي العراق نظمت الجمعية العراقية للعلوم التربوية والنفسية في آذار عام 1980 حلقة دراسية حول استخدام التقويم في المرحلة الثانوية تمت خلالها مناقشة عدة بحوث علمية حول استخدام الوسائل التقويمية في العملية التعليمية والبدائل المقترحة لتطويرها كي تحقق مستوى أفضل في الميدان التربوي بهدف الوصول إلى تقويم موضوعي عادل لمستوى الطلبة.

ومما تحقق في مجال التقويم في البلاد العربية ما تقوم به بعض تلك البلدان من استخدام بعض الخبراء في مجال التقويم والقياس بالتعاون مع منظمة اليونسكو وإنشاء مكاتب أو مؤسسات رسمية تابعة لوزارات التربية تتولى الإشراف على تنظيم عملية التقويم في المؤسسات التعليمية. كما ترسل الدول العربية العديد من البعثات والإجازات الدراسية إلى الدول المتقدمة للحصول على شهادات عليا في مجال التقويم والقياس التربوي وما ينتج عن ذلك من تطور في هذا الميدان عند عودة هؤلاء وما يقدمونه من أعمال تؤثر إيجابياً على نقل ما استجد من معطيات ونظريات حديثة في ميدان التقويم إلى بلدانهم.

تجربة سلطنة عُمان:

شهدت نهاية القرن العشرين تطورات أكبر في مجال التقويم والقياس لاسيما بعد أن استخدمت التقنيات التربوية المتقدمة والحاسوب في مجالات عديدة ومنها المجال التربوي حيث أصبحت عمليات تحليل البيانات واستخدام الأساليب الإحصائية المتقدمة تجرى عن طريق الحاسوب رغم كبر حجم العينات المستخدمة، مما أضاف عنصرا جديدا تمثل في إحداث نقلة نوعية في نظام تقويم تعلم الطلاب في مدارس العديد من الدول العربية كما تبنت بعض تلك الدول ومنها سلطنة عُمان - على سبيل المثال- نظاماً جديداً يتسم بالتنوع والاستمرارية والشمول يقوم على استخدام طرق وأدوات تقويمية متنوعة من بينها (نظام التقويم التكويني المستمر) في مدارس التعليم الأساسي ومجموعة من مدارس التعليم العام المطبقة لمشروع تطوير تقويم أداء الطلاب. وتتم عملية التقويم التكويني المستمر في جميع المواد الدراسية، وتحسب العلامة الكلية للطالب في بعض المواد الدراسية (التربية الإسلامية، اللغة العربية، اللغة الإنجليزية، الدراسات الإجتماعية، العلوم، والرياضيات، المهارات الحياتية) وفق النسب الآتية:

- (40%) للاختبارات القصيرة بما لا يقل عن اختبارين في الفترة الواحدة (حيث قسمت السنة الدراسية إلى (أربع فترات).
- (60%) لأدوات التقويم التكويني المستمر. أما المواد المتبقية مثل (الرياضة المدرسية، المهارات الموسيقية، الفنون التشكيلية، تقنية المعلومات) فتكون النسب كالآتي:
- (20%) للاختبارات القصيرة.
- (80%) لأدوات التقويم التكويني المستمر.

ولضمان نجاح هذا الأسلوب من التقويم وتوثيق نتائجه تمَّ تشكيل (لجنة متابعة التحصيل الدراسي) في كل مدرسة لمتابعة التلاميذ ضعاف التحصيل وذوي صعوبات التعلم، والتلاميذ المتفوقين والموهوبين واقتراح البرامج

العلاجية والإثرائية المناسبة لهم. يضاف إلى ذلك إجراء آخر يقوم خلاله كل معلم بإعداد ملخص إنجاز لكل تلميذ في نهاية العام الدراسي يبين فيه جوانب القوة والضعف في إنجاز التلميذ لكل مادة وتوضح هذه الإنجازات ما يقدمه الطالب من أعمال كتابية أو فنية متنوعة ومدى إنجازه للأهداف المحددة في ملف واحد لكل مجال أو مادة يسمى ملف أعمال التلميذ أو الملف الوثائقي Portfo-lio) يطلع عليه ولي أمر الطالب في نهاية كل فترة دراسية مع تقارير الأداء وملاحظات المعلم الوصفية حول بعض الأعمال. كما تعد المدرسة ملفاً لكل تلميذ (الملف التراكمي) في نهاية العام الدراسي يتضمن (ملخصاً لإنجاز التلميذ في كل مادة، تقرير الأداء، المشاركات المتميزة للتلميذ، البيانات الثبوتية، تقرير الأخصائي الاجتماعي أو الطبيب إذا كان التلميذ يعاني من مشكلة اجتماعية أو صحية والبطاقة الشخصية للطالب المعدة من قبل الوزارة، وأخيراً تقرير لجنة التحصيل الدراسي (إن كان التلميذ تحت المتابعة). أما التقديرات الكلية النهائية لتحديد مستوى تحصيل الطالب فتكون على النحو الآتي:

الدرجة	المؤشر	التقدير
100-90	أ	ممتاز
89-80	ب	جيد جداً
79-65	ج	جيد
64-50	د	مقبول
49 فأقل	هـ	يحتاج إلى متابعة

وبما أن الدرجة الصغرى التي تمثل أدنى مستوىً من الأداء هي (50) فإن الدرجة (49) تمثل مستوى الأداء الأقل من المستوى الأدنى وفي هذه الحالة لا يبقى التلميذ راسباً في صفه بل ينقل إلى صف أعلى مع تقديم البرامج العلاجية له في ذلك الصف. أما إذا حصل هلى تقدير (هـ) في مادتين أو

أكثر فإن الأمر متروك إلى لجنة متابعة التحصيل الدراسي لتقرر إنتقاله إلى الصف الأعلى أو إعادته في صفه مع وضع خطة علاجية تتناول جوانب الضعف وأسبابه وخطة العلاج المقترحة. (وزارة التربية والتعليم - دائرة التقويم التربوي، ٢٠٠٨).

أما بالنسبة لطلاب المراحل الأخرى التي تبدأ من الصف الخامس وحتى الصف العاشر (تعليم أساسي أو تعليم عام) فوضع لها نظام تقويم خاص يختلف حسب المواد الدراسية التي قسمت إلى مجاميع خصصت فيها للاختبارات القصيرة (60%) ولأدوات التقويم التكويني المستمر (40%) فيما عدا المواد الرياضية والفنية خصصت لها (20%) للاختبارات القصيرة، و(80) لأدوات التقويم التكويني المستمر. أما في الصف الثاني عشر وهو الذي يمثل نهاية المراحل الدراسية قبل المرحلة الجامعية يتم تقويم أداء الطالب على النحو الآتي:

أولاً: مواد المجموعة (أ) ما عدا (الرياضة المدرسية والفنون التشكيلية) يتم تقويمها كالآتي:

- (60%) للاختبارات القصيرة والفصلية منها (15 درجة) لكل من الاختبارين القصيرين و(30 درجة) للاختبار الفصلي.
- (40%) أدوات التقويم الأخرى.

وقد تم تحديد خصائص الاختبارات القصيرة بأنها كتابية تؤدى بصورة فردية في وقت لا يتجاوز (عشرين دقيقة) لكل اختبار. أما الاختبار الفصلي فيتم إعداده وفق المواصفات الاختبارية الخاصة بكل مادة دراسية وتقوم المدرسة بالإشراف على مستلزماته المادية من طباعة وإشراف وتصحيح وإرسالها إلى المنطقة التعليمية التي تنتمي إليها المدرسة.

أما بالنسبة لمواد تنمية المهارات الفردية (الرياضة المدرسية والفنون التشكيلية فيتم تقويمها تقويماً مستمراً على النحو الآتي:

- (20%) للاختبارات القصيرة و (80%) لأدوات التقويم المستمر. وهناك مواد دراسية تسمى مواد المجموعة (ب) يتم توزيع درجاتها على النحو الآتي:
- (70%) لامتحانات نهاية الفصل الدراسي التي يتم إعدادها وتصحيحها وإعلان نتائجها من قبل الوزارة.
- (30%) للتقويم التكويني المستمر.

ولقد جاء اهتمام وزارة التربية العمانية بالتقويم التكويني المستمر كأسلوب يهدف لتحسين العملية التعليمية من خلال التعرف على نواحي القوة والضعف ومدى تحقق الأهداف التربوية والاستفادة من التغذية الراجعة في تعديل المسار نحو تحقيق تلك الأهداف، وهذا الأسلوب يعتمد على الملاحظات اليومية، والأنشطة الصفية وغير الصفية والاختبارات القصيرة والفترية وهو تقويمٌ مستمرٌ ملازمٌ للعملية التعليمية من بدايتها وحتى نهايتها. (وزارة التربية والتعليم - دائرة التقويم التربوي، 2004). وما من شك بأن تعدد وسائل التقويم يؤدي إلى حكم أفضل لمستوى تحصيل الطالب، وهذا التوجه يتماشى مع توصيات الدارسات والمختصين في مجال القياس والتقويم مثل (Kim and Sunderman, 2005 Perkins, 2005, Darling, 2006, Nitko, 2006, lee, 2007) ولم تقتصر جهود التغيير والتطوير المتبعة في النظام التربوي لسلطنة عُمان على الاهتمام بتقويم أداء الطالب بل إمتدت إلى تقويم الأداء الإداري والمدرسي المؤسسي، من خلال تطبيق (مشروع تقويم الأداء المدرسي) الذي يركز على ثلاثة عناصر هي (التعلم والتعليم والإدارة المدرسية) باستخدام أدوات تقويم متنوعة كالملاحظات والمقابلات والاستبيانات والتحليل وغيرها، من أجل جمع أكبر ما يمكن من الأدلة للحكم على العناصر الثلاثة المذكورة آنفاً وفق المعايير التي توضع لكل عنصر.

وتُشير أدبيات وزارة التربية إلى أن هذا المشروع الذي يعتمد على تشخيص الوضع الحالي للأداء الفني والإداري في مدارس السلطنة من خلال

إبراز جوانب القوة والضعف وتحديد الأولويات يهدف لتحقيق جملة من الأهداف لعل أبرزها، وضع نظام شامل لتقويم الأداء المدرسي بمجالاته المختلفة، وإعداد فرق مؤهلة للقيام بعملية تقويم الأداء المدرسي ومساعدة المدرسة على تطوير نظام ذاتي وفاعل لتقويم أدائها وتطويره، والتعرف على مدى تحقق أهداف سياسة التعليم بالسلطنة من خلال معرفة ما حققته المؤسسات التعليمية واستخدام نتائج التقويم في وضع وتطوير السياسات التعليمية (وزارة التربية والتعليم - دائرة تطوير الأداء المدرسي، 2004)، وسلطنة عُمان بذلك تنتهج نهج العديد من دول العالم التي تسعى لتطوير إداء مؤسساتها التعليمية.

ومن أجل تنفيذ عملية تقويم الأداء المدرسي وفق أسس علمية سليمة شكلت الوزارة لجان عدة منها لجان تقويم من الكادر المدرسي ويقوم بعملية التقويم الذاتي Self Evaluation، ولجان أخرى يشارك فيها الجهاز الإداري والإشرافي التابع للوزارة ويقوم بعملية التقويم الخارجي External Evaluation وتم اختيار ثمان مدارس من كل منطقة تعليمية عدا مسندم والوسطى تم اختيار أربعة مدارس لكل منها وبدأ بتطبيق هذا المشروع تجريبيا اعتباراً من السنة الدراسية 2004/2003 ولازال مستمراً. ولا شك أن نماذج هذين النوعين من التقويم تؤديان وظيفة كبيرة لترصين النظام التعليمي، ولابد من القول أن السلطنة تعتبر الرائدة الأولى في هذا الجانب بالنسبة للأقطار العربية الأخرى.

مراجع الفصل الأول

1- الإمام، مصطفى عبد الرحمن، أنور حسين وصباح حسين العجيلي (1995). **التقويم والقياس**. بغداد: دار الحكمة للطباعة والنشر.

2 - الزويعي، عبد الجليل، بكر، محمد إلياس، والكناني إبراهيم، (1987). **الاختبارات والمقاييس النفسية**. الموصل: مطبعة جامعة الموصل.

3- وزارة التربية والتعليم، دائرة تطوير الأداء التربوي (2004). **المذكرة التفسيرية لمستجدات التقويم التربوي بمدارس التعليم الأساسي والتعليم العام**. مسقط.

4- وزارة التربية والتعليم، دائرة التقويم التربوي (2008). **الوثيقة العامة لتقويم تعلم الطلبة للصفوف** (1-12). مسقط..

5- لندفل، س. م (1968). **أساليب الاختبار والتقويم في التربية والتعليم**، ترجمة عبد الملك الناشف وسعيد التل. بيروت: المؤسسة الوطنية للطباعة والنشر.

6- Guba, E. G., & Lincoln Y. S. (1981). **Effective Evaluation**. California: Jossey- Bass.

7- Kline,P.(2000). **Hand book of Psychological Testing (2nd ed).** New York: Routledge.

8- Kubizyn, T., & Borich G. (1987). **Educational Testing and Measurement: Classroom Application and Practice. (2nd ed).** Illinois: Scott, Foresman and Company.

9- Rossi P. H., & Freeman H. E. (1982). **Evaluation Asystematic Approach**, (2nd ed), California: SAG publications, Inc.

10- Stufflebeam, D. L., & Shinkfield A. J. (1985). **Systematic Evaluation.** Boston: kluwer- Nij off Publishing.

11- Worthen B. R., & Sanders J. R (1973). **Educational Evaluation: Theory and Practice**. California: Wadsworth Publishing Company, Inc.

12- Worthen, B.R,. & Sanders, J.R. (1987). **Educational Evaluation: Alternative Approaches and Practical Guidelines**. New York: Longman.

الفصل الثاني

مفاهيم أساسية

التقويم

أهداف التقويم

مجالات التقويم

أنواع التقويم : الكمي - النوعي - البنائي - الختامي - المكبر - المصغر - الخارجي - الذاتي

فلسفة ومبادئ التقويم الذاتي

أهداف التقويم الذاتي

خصائص التقويم الذاتي الفعال

التخطيط للتقويم الذاتي

العناصر التي يشملها التقويم الذاتي

مراحل وإجراءات التقويم ا لذاتي

كتابة التقرير الختامي

أنواع تقارير التقويم

2

الفصل الثاني

مفاهيم أساسية

التقويم Evaluation

لقد أعطيت للتقويم تعريفات عدة يتشابه بعضها مع البعض، وتختلف عن بعضها شيئاً ما في الهدف أو الشمول أو الاستخدام لهذا المفهوم الذي أصبح اليوم عملية تثير الكثير من الاهتمام والجدل في نفس الوقت، حتى قال عنه بعض المهتمين به: أنه الموضوع الذي يُناقش بكثرة ويستخدم بقلة في بعض الأنظمة التربوية (Worthen and Sanders, 1973).

إن أول ما يتبادر للقارئ العادي وهو يسمع كلمة (التقويم) بأنها تعني (التعديل)، أي تعديل اعوجاج الشيء أو إصلاحه، بينما يردد البعض أن التقويم يعني الاختبارات، وبالدقة (الامتحانات)، وهناك من يستخدم كلمة (القياس) ليقصد بها (التقويم).

وعبر هذا الخلط والاستخدام الخاطئ تبرز التعريفات التي تمثل مدارس فكرية أو آراء شخصية ناشئة عن خبرة ميدانية من بعض العاملين في مجال التقويم. وسنستعرض في السطور القادمة بعضاً من تلك التعريفات التي قيلت في التقويم ونستطلع أوجه التشابه والاختلاف بينها لنصل إلى الصورة الواضحة وبأبسط حالاتها قدر الإمكان.

يقال في اللغة العربية: قوّمت السلعة بمعنى ثمنت. ويقال أيضاً: قوّم الشيء أي قدّر قيمته ووزنه، وقوّم الأستاذ أعمال طلابه أي أعطاها قيمة ووزناً.

ويؤكد المعنى اللغوي للتقويم على أنه التثمين، أي إعطاء وزن وقيمة للشيء المقوّم. ولاشك أن إعطاء القيمة والوزن في النظرة الحديثة لمفهوم التقويم تعني جزءاً من هذه العملية وليست العملية كلها. وتتفق هذه التوضيحات مع ما حددته (لجنة المعايير الخاصة بالتقويم التربوي The Joint

(Committee on Standards for Educational Evaluation) من تعريف للتقويم بأنه:

"التحديد المنظم لقيمة أو جدارة الموضوع المقوّم". ويلاحظ أن هذا التعريف والتعريفات السابقة تؤكد على "القيمة" Value، والتي تعرّف في علم النفس بأنها:

"أي مظهر لموقف أو حدث أو شيء معين يكون محاطاً برغبة تفضيلية، حسن، رديء، أمر مرغوب، غير مرغوب، أو ما شابه ذلك". لكن ما يميز التعريفات الحديثة لمفهوم التقويم عن التعريفات السابقة الذكر أنه لا يكتفي بإعطاء قيمة أو وزناً للشيء المقوّم، بل يمتد إلى أبعد من ذلك فيضيف عنصراً آخر ألا وهو إصدار الأحكام Judgments حول الشيء المقوّم. ويؤكد أصحاب هذا الرأي على أن أية دراسة تقويمية لا تشير إلى جودة أو رداءة الشيء المقوّم - أي الحكم عليه- لا يمكن أن يطلق عليها كلمة تقويم. (Worthen and Sanders, 1973)

ومن التعريفات التي أكدت على جانب الحكم ما أورده (وايلز Wiles) عندما وصف التقويم بأنه: عملية تصدر عنها الأحكام التي تُستخدم كأساس للتخطيط. ويدعم تنبرنك (Tenberink) ذلك بقوله: أن التقويم عملية الحصول على المعلومات اللازمة للوصول إلى الأحكام التي بواسطتها يمكن اتخاذ القرارات.

ويتفق المؤلفون مع ما تشير إليه التعريفات السابقة من ضرورة امتداد التقويم إلى إصدار أحكام، إذ أن هذه العملية إذا ما خلت من عمل أو إصدار الأحكام تصبح عملية وصف مجرد Description، ووصف الشيء يختلف عن تقويمه لأن الوصف يعني استعراض خصائص أو سمات الشيء كما يراها الواصف. أما إذا أضفنا إلى (الوصف) عملية (الحكم) فإننا نبرز قيمة الشيء. ولابد لنا هنا أن نذكر عنصراً ثالثاً يضاف إلى عنصري (الوصف) و (الحكم)، ذلك هو (المعيار Standard) الذي نُقارن به خصائص أو سمات الشيء المقوم قبل إصدار الحكم عليه. والمعيار أما أن يكون ضمنياً أو بارزاً، أي إما أن يكون شاخصاً في ذهن المقوّم أو مادياً كمياً محدداً يُسهّل عملية المقارنة

وإصدار الأحكام المستندة إلى أرقام أو كميات. ويؤيد ضرورة وجود المعايير العديد من العاملين في مجال التقويم عندما يؤكد بعضهم على أن هدف التقويم هو تحديد الوضع الراهن للشيء المقوّم بمقارنة هذا الوضع أو الحالة بمجموعة من المعايير أو المحكات.

وبناءً على ما تقدم يمكن القول أن العناصر التي تتكون منها عملية التقويم تشمل: الوصف Description والمعيار Standard والحكم Judgmant وبذلك يمكن تقديم تعريف مبسط للتقويم بأنه: عملية وصف الشيء والحكم عليه وفقاً لمعيارٍ معين.

ويتضمن الوصف جمع بيانات ومعلومات كمية أو نوعية عن الشيء المقوّم سواء باستخدام وسائل قياس أو باستخدام الملاحظة والخبرة الشخصية، وبعد ذلك تتم مقارنة هذه المعلومات بمعيار معين. والمعايير مختلفة ومتنوعة فقد يكون المعيار هو الهدف الذي نسعى إلى تحقيقه في برنامج معين، وقد تكون المعايير أوزان أو قياسات معينة ثابتة ومجربة ومحددة. وبعد عملية المقارنة هذه نبدأ بإصدار الأحكام التي تتضمن الإشارة إلى الجوانب الإيجابية و السلبية في الموضوع المقوّم، وقد تُقدم الأحكام أو النتائج التي تمخض عنها التقويم إلى صانع القرار Decision Maker وتمهد له في اتخاذ قرارات معينة استناداً إلى ما يقدمه المقوم من معلومات وإثباتات مسندة علمياً أو منطقياً ومدعمة بالواقع. كما أن المقوّم ومن خلال المعرفة التي تكونت لديه من معايشته للبرنامج المقوّم، وما تجمعت لديه من معلومات وبيانات تم تحليلها من قبله سيكون قادراً على وضع اقتراحات توفر لصانع القرار أو الشخص المسؤول استبصارات تجعل ذلك الشخص في حالة تمكنه من اتخاذ قرارات صائبة.

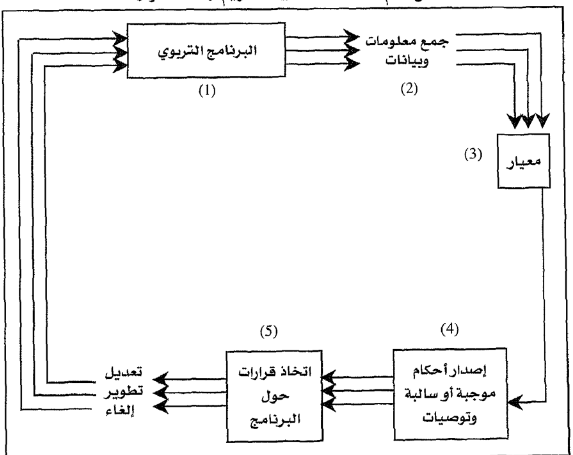
إن عملية اتخاذ القرارات تعتمد- بلا شك- اعتماداً كلياً على جودة ودقة المعلومات التي تُقدم من قبل المقوّم بشأن البرنامج التربوي أو أية فعالية أو نشاط.

ومما يجدر ذكره هنا أن متخذ القرار في كثير من الأحيان ليس هو المقوّم نفسه وإنما هو صاحب الصلاحية في إجراء التغيرات في مؤسسة ما.

وهناك بعض الحالات يكون فيها المقوّم صانع القرار نفسه مثال ذلك المدرس الذي يقوّم مستوى طلبته للوقوف على مدى استيعابهم ودرجة تحصيلهم في مادة ما، لذلك فهو يستطيع اتخاذ قرار بشأنهم كقراره بنقلهم إلى صف أعلى أو إبقاء البعض منهم في نفس الصف.

لقد تَحَدثَ الكثير من المهتمين بالتقويم عن علاقة عملية اتخاذ القرار بنتائج التقويم، وتطرقوا إلى هذه العلاقة في التعريفات التي قدموها عن التقويم. إذ يعرف (تنبرنك) التقويم بأنه عملية الوصف الدقيق للحصول على المعلومات المفيدة للحكم على بدائل القرارات. ويؤكد ستفلبيم (Stuffelbeam) على أن التقويم عملية تخطيط تهدف إلى الوصول إلى جمع بيانات مفيدة تساعد متخذ القرار على الاختيار من بين بدائل متعددة. ويمكن تمثيل العلاقة بين التقويم واتخاذ القرارات بالمخطط الأتي:

الشكل رقم (1-2) العلاقة بين التقويم واتخاذ القرارات



ويرى كل من جوبا وستيفلبيم (Guba and Stuffelbeam) نظراً للدور الأساسي للمقوّم المتمثل في توفير معلومات لصانع القرار، فإنه (أي المقوّم) يجب أن يكون متخصصاً في تشخيص وجمع وتحليل وكتابة المعلومات. أما ستيك (Stake, 1967) فيقترح أن المقوّم هو الذي يجب أن يعمل الأحكام ويطرح بدائل القرار التي إذا ما صيغت بدقة فإن أي شخص مسؤول سيكون قادراً على الوصول إلى استنتاجات لاختيار الأفضل منها.

ومما تقدم يلاحظ القارئ ان تعريفات التقويم لا تختلف عن بعضها كثيراً رغم التباين الظاهري في وجهات النظر وفقاً للوظائف التي يشغلها القائمون بالتقويم في المؤسسات التربوية حيث يعرّف المعلمون التقويم بعبارات وصفية أو كمية تتعلق بالمواد التي يدرسونها، والمرشدون التربويون يعرفونه بقدر تعلقه بتخصصات الطلبة وميولهم المهنية، ومدراء المدارس يعرفونه في ضوء سلوكيات الطلبة، أما الآباء فربما يهمهم في التقويم مدى ما يوفره لهم من مؤشرات تتعلق بمستوى تحصيل ابنائهم، بينما يهتم مقومو البرامج التربوية بجمع المعلومات عن البرنامج ومقارنتها بما تحقق من اهداف ذلك البرنامج لأجل إصدار أحكام حولها (علام، 2003).

أهداف التقويم:

من خلال التعريفات السابقة أصبح بالإمكان تحديد الأهداف أو الأغراض التي يمكن تحقيقها بواسطة عملية التقويم ومنها:

1- وصف وإصدار أحكام عن البرامج والنشاطات التربوية بطريقة منظمة ومدروسة.
2- تقرير مدى ما تحقق من أهداف برنامج أو أسلوب تربوي معين وما لم يتحقق بعد تطبيقه لفترة زمنية معينة.
3- تحديد نواحي العجز أو القصور من أجل التوصل إلى وضع خطة لتلافيها بالأساليب والمعالجات المناسبة.
4- تحديد الخصائص الإيجابية للبرنامج أو النشاط التربوي والتي تلبي الحاجات الاجتماعية أو الأهداف التربوية العامة والخاصة لتعزيزها وضمان استمرارها.

5- تقرير استمرارية أو تطوير البرنامج التربوي أو إيقافه.
6- توفير معلومات مفيدة وضرورية لاتخاذ قرارات صائبة تتعلق بالنشاط والبرنامج أو العنصر الذي يتم تقويمه (الطالب، المعلم، المنهج، الأهداف... الخ).

مجالات التقويم:

بما أن التقويم يمثل جانباً مهماً من جوانب العملية التربوية فلا بد له أن يكون شاملاً ومستمراً لكافة مجالاتها ومنها:

1) تحديد تحصيل الطالب:

من الواضح أن عملية تحديد مستوى تحصيل الطالب في المواد الدراسية تعتبر من أولى المهام التي يقوم بها المدرس. ونظراً لأن ما هو مطلوب من المدرس هو إحداث تغيرات مرغوبة في سلوك الطالب فإنه (المدرس) يحتاج إلى ما يساعده على التأكد من تحقيق هذا الغرض، أي قياس ما تعلمه الطالب. ولذلك فإن بإمكانه استخدام ما يناسبه من أدوات القياس المعروفة في المجال التربوي ومنها:

أ. **الأساليب غير الاختبارية:**

أي أن بإمكان المعلم التعرف على مدى ما تعلمه الطالب دون الحاجة إلى تطبيق اختبارات معينة، وذلك من خلال ملاحظة سلوك الطالب في الصف وخارجه، مثل إجاباته ومناقشاته وتحضيره اليومي وإنجاز واجباته اليومية المكلف بها على الوجه الأكمل. ويمكن للمعلمين استخدام هذا الأسلوب بجانب استخدام الاختبارات التحصيلية، ليكوّنوا صورة أفضل وأشمل عن مستويات طلبتهم من خلال تنوع أساليب التقويم.

ب. **الاختبارات التي يصممها المعلم:**

وتعتبر من الوسائل القديمة الاستخدام والشائعة في المجال التربوي لتحديد أو تقويم تحصيل الطالب. وعلى الرغم من أن بعض المعلمين يمكنهم عمل أنواع ممتازة من الأسئلة الامتحانية نتيجة لخبرتهم وطول تجربتهم في المهنة، إلا

أن هناك البعض الآخر ممن تنقصهم الخبرة ويأتون بأسئلة كثيراً ما تكون مربكة ولا تُقيس إلا جوانب معرفية تعتمد على حفظ المادة المقروءة مما ينتج عنها أضراراً أكثر من الفائدة المرجوة. ولذلك سنركز في فصول قادمة من الكتاب على موضوع صياغة الأسئلة الامتحانية (التقليدية منها والموضوعية) ومجالات استخدامها وأنواعها، لكي تكون عوناً للمعلمين والمدرسين وأساتذة الجامعات في هذا المجال.

ج. **اختبارات التحصيل المنشورة:**

يُشيع استخدام هذا النوع من الاختبارات التحصيلية في الدول المتقدمة حيث توجد مؤسسات خاصة لبناء وتطوير هذه الاختبارات تحت إشراف خبراء ومتخصصين في موضوعات دراسية مختلفة ولمراحل دراسية متنوعة، ويمكن للمعلم اقتناءها واستخدامها لتقويم تحصيل طلبته في مادة ما. ويتطلب استخدام هذه الاختبارات إطلاع المعلم على أنواعها والأساليب الإحصائية الخاصة بتحليل نتائجها وأن تكون هذه الأنواع مألوفة لديه لكي يميز بينها لاختيار ما يناسبه منها، وتسمى هذه الاختبارات "بالاختبارات المقننة -Stan dardized Tests" أو الاختبارات المعيارية والتي تطبق على عينات كبيرة من الطلبة لغرض التحقق من صدقها وثباتها وتحليل فقراتها وتحديد معايير خاصة لتفسير علامات الطلبة الممتحنين بها وتوضيح ظروف وقواعد إجرائها.

2) تحديد استعدادات وقدرات الطلبة:

تعتبر الاستعدادات والقدرات التي يملكها الطالب من الأمور الأساسية التي يُبنى عليها التعلم اللاحق، ولذلك يوصي المربون بضرورة الكشف عن مثل هذه القدرات وتحديدها قبل البدء بإعطاء خبرات جديدة للطالب أو لتنمية هذه القدرات في الاتجاه السليم والمناسب. وتستخدم للكشف عن الاستعدادات والقدرات أساليب تقويمية شتى، إلا أن المعلم يمكن أن يتعرف على تلك القدرات حتى من خلال الملاحظة والمناقشة حيث يساعده ذلك على وضع خطة للبناء عليها واستغلالها. مثال ذلك ما يقوم به المعلمون من التعرف

على استعدادات وخبرات طلبتهم في مادة الرياضيات للصف الرابع - على سبيل المثال - قبل البدء بإعطاء موضوعات أعلى مستوى في الصف الخامس.

وعلى الرغم من جدوى مثل هذه الأساليب البسيطة في الكشف عن قدرات الطلبة في مجالات شتى تبقى هذه الأساليب يعوزها الضبط والدقة والموضوعية، لذلك يضع المتخصصون بأساليب القياس التربوي اختبارات خاصة لتحديد قدرات الطلبة وهي ما تسمى (اختبارات الاستعداد Aptitude Tests) وهي متنوعة وتستخدم بكثرة في البلدان المتقدمة لهذا الغرض نذكر للقارئ فيما يلي بعضاً منها للتعريف:

- اختبارات الاستعداد المدرسي General Scholastic Aptitude Tests.
- اختبار الاستعداد لمواد دراسية محددة. Test of Aptitude for Specific Subjects

3) تحديد رغبات الطالب:

من أجل الوصول إلى تخطيط ناجح في المجالين التربوي والمهني لمستقبل الطالب أصبح من الضروري للوالدين والمعلمين أن يكونوا على إطلاع كاف ببعض التقديرات غير الرسمية (النظامية) لرغبات الطلبة. ومع ذلك فإن هناك عدداً من الاختبارات والقوائم التي توفر أساليب موضوعية لتحديد تلك الرغبات والأمور المفضلة عنده من أجل مساعدته وإرشاده نحو الأعمال أو التخصص الدراسي الذي يتماشى مع تلك الرغبات.

4) تحديد شخصية الطالب:

مثلما يهتم ميدان التقويم التربوي بموضوع المستوى المعرفي للطالب فإنه يشمل تحديد الجوانب الإيجابية والسلبية لشخصية الطالب. ويجري هذا التقويم عبر وسائل واختبارات شخصية ونفسية كما يمكن أن تستخدم الملاحظة البسيطة من قبل المعلم لتحديد بعض الأنماط السلوكية لشخصية الطالب. ولا ينصح المختصون بعلم النفس قيام معلم الصف باستخدام الاختبارات الخاصة

بالشخصية بل يمكن استخدامها من قبل أخصائيين في علم النفس والإرشاد النفسي لأنها تتطلب مستوىً معيناً من الخبرة والمعرفة لتطبيقها وتفسير نتائجها.

ومن الأمور الأخرى المتعلقة بهذا الشأن أن التقويم يهتم بتحديد بعض جوانب الشخصية والتركيز على الجوانب التي تحتاج إلى مساعدة حيث يبرز دور الأب والمعلم والمرشد في تقديم المعونة حسب الإمكانية للطالب بعد تشخيص تلك الجوانب.

5) تشخيص صعوبات المتعلم:

يستفيد المعلم من عملية التقويم للتعرف على بعض التلاميذ الذين يتخلفون عن الآخرين من زملائهم للوصول إلى المستوى المناسب في التحصيل. ويساعد التقويم في هذا المجال توجيهه المعلم للتعرف على العوامل التي قد تؤثر على تحصيل هؤلاء وكيفية وضع خطة مناسبة لمعالجتهم ورفع مستواهم وأخذ تلك العوامل بنظر الاعتبار ومحاولة تخفيفها أو إزالتها قدر الإمكان.

6) التعرف على الجوانب التي تعرقل عملية التعلم ومعالجتها:

يفيد التقويم في هذا المجال في تسليط الضوء على العوامل التي تؤثر سلباً على تعلم الطلبة. فإذا ما أجرى المعلم اختباراً لطلبته في مادة ما وتوصل - على سبيل المثال- إلى أن 80% رسب من مجموع طلبته، فإنه حريٌ به أن يسأل نفسه أسئلة عديدة منها: هل أن السبب في نسبة الرسوب العالية هذه يعود إلى الكتاب المستخدم في التدريس بسبب عدم مراعاته لمستويات الطلبة العمرية أو العقلية؟ أم هل يعود السبب إلى أمور تتعلق بحياة الطلبة وظروفهم العائلية خارج المدرسة؟ بل قد يكون السبب في طريقة التدريس التي استخدمها المدرس؟ أو أن صياغة الأسئلة الامتحانية تتسم بالغموض والتشعب وعدم الوضوح بحيث لم يستطع الطلبة فهم المقصود منها؟ كل هذه التساؤلات وأخرى غيرها يمكن أن تدفع المقوّم أو مدرس الطلبة إلى وضع يده على العوامل الأساسية أو على أيها أكثر تأثيراً لكي يبدأ بوضع خطط للمعالجة وبذلك يمكن أن يساهم في التخفيف أو إزالة المعوقات التي تعرقل عملية التعلم.

7) الإرشاد التربوي والنفسي:

مما لا شك فيه أن الإرشاد التربوي والنفسي يمثل جانباً مهماً في العملية التربوية لأنه يُبصر المرشدين بتحديد الاضطرابات والمشكلات التي يعاني منها الطالب كصعوبة إقامة علاقات موفقة مع الزملاء، أو عدم التكيف مع المحيط المدرسي بصورة عامة مما يولد صراعات واضطرابات نفسية تؤدي إلى كرهه للدراسة أو المحيط المدرسي. وهنا يقوم المرشد التربوي بالاستفادة من نتائج تطبيق بعض المقاييس النفسية لتشخيص الحالة التي يعاني منها الطالب ويضع الخطط اللازمة لمعالجتها واعادة الطالب الى الحياة السوية وتشجيعه على الانسجام مع المحيط الذي يعيش فيه. أما فيما يخص التوجيه التربوي والمهني فإن المرشد المتخصص يستطيع استخدام اختبارات الاستعداد والتحصيل والشخصية، ويجري المقابلات المباشرة مع الطلبة فردية كانت أم جمعية لغرض توفير المعلومات عن الطالب حيث تساعده هذه المقابلات والمعلومات في توجيه الطلبة إلى الدراسة أو المهنة التي تناسب قدرات ورغبات كل منهم لضمان نجاحهم في تلك المهن أو التخصصات الدراسية التي يتقدمون للقبول فيها.

8) اختيار الأهداف التعليمية:

للتقويم دور مهم في هذا المجال، ونظراً لأن التربية تهدف إلى إحداث تغيرات مرغوبة في سلوك الطلبة من خلال التعلم والتعليم فإن تقويم هذه الأهداف أو وسائل تحقيقها يساعد المؤسسات التعليمية والقائمين على شؤون التربية على فهمٍ واضح للغايات والأهداف التي حددت مسبقاً وما إذا نجحت فعلاً في إحداث التغيرات المطلوبة عن طريق المادة التعليمية وطرق التدريس أم لا، مما يستدعي تعديل أو تغيير تلك الأهداف بموجب نتائج التقويم وتعديل أو تكييف الوسائل التي تستخدم في تحقيقها. لذلك فإن التقويم على الرغم من أنه يستخدم في كثير من الأحيان الأهداف كمعيار للحكم على ما يتم تحقيقه في عملٍ ما، فإن المقوّم قد يجد في بعض الأحيان أن هناك ضرورة لتعديل أو تغيير الأهداف المحددة للبرنامج.

9) تصنيف الطلبة:

في كثير من الأحيان يتطلب الأمر تصنيف الطلبة في الصف وفقاً لمتغيرات معينة توجب استخدام وسائل القياس والتقويم التي تحقق هذه الغاية مثال ذلك: استخدام اختبارات الذكاء واختبارات الاستعدادات والقدرات والميول في التعرف على سمات الطلبة وتصنيفهم إلى مجاميع ومن ثم التعامل معهم وفقاً لتلك السمات وتنميتها إلى أعلى حد ممكن. ويجري استخدام هذا الأسلوب في البلدان المتقدمة لتحديد الطلبة الموهوبين أو المتخلفين من أجل وضع برامج تربوية ونفسية خاصة تتناسب وقدراتهم ومستوياتهم والأخذ بأيديهم (لاسيما أصحاب الاحتياجات الخاصة) للوصول إلى حالة أفضل سواء بوضعهم في صفوف خاصة أو رعايتهم في نفس صفوف أقرانهم الأسوياء. كما يمكن تصنيف الطلبة حسب رغباتهم وهواياتهم لمحاولة إشباعها وتطويرها من قبل المعلم.

10) تقويم كفاءة المدرس:

قد تبدو معظم أغراض التقويم التي سبق الحديث عنها متعلقة بالطالب، ولكن ذلك لا يعني أن التقويم يركز على الطالب وحده بل يمتد إلى تقويم المدرس أو المعلم للتعرف على كفاءته ومستواه، سواء كان ذلك لأغراض تتعلق باتخاذ قرارات إدارية، أو لأغراض تطويرية تستهدف رفع مستواه العلمي والتربوي من خلال برامج التدريب أثناء الخدمة.

لقد تعددت الوسائل والأساليب المتبعة لتقويم كفاءة وفاعلية المدرس منها على سبيل المثال: الزيارات الصفية التي يقوم بها المشرف أو مدير المدرسة للمعلم. واستخدام مقاييس ثابتة وصادقة لتقويم كفاءته من قبل المسؤولين الإداريين وهو ما يطلق عليه (التقويم الإداري Administrative Evaluation) أو استخدام تقويمات وتقديرات الطلبة الذين يدرّسهم المدرس وهو ما يسمى (تقديرات الطلبة Student Ratings) كما أن هناك ما يسمى (تقويم الزملاء

Peer Evaluation) عندما يطلب رأي زملاء المعلم من الذين يعملون معه في نفس المدرسة أو الكلية عن مستواه العلمي أو التربوي وحتى الشخصي. كما أن هناك أسلوباً آخر يمارسه المدرس لتقويم نفسه بنفسه وهو ما يطلق عليه (التقويم الذاتي Self Evaluation) ويمكن استخدام استمارة لهذا الغرض تحدد فيها نقاط عديدة لشخصية المعلم وواجباته التربوية والعلمية وعليه أن يؤشر أمام كل خاصية الدرجة التي يستحقها هو نفسه ويصل إلى نتيجة تمثل تقويمه لنفسه بموجب هذا المقياس. إن هذه الوسائل التي أشير إليها تهدف إلى إعطاء صورة معينة عن المدرس يستطيع الاستفادة من نتائجها في تطوير شخصيته وعمله في التدريس.

11) إعادة النظر في المناهج:

إن الاستمرار باستخدام وسيلة تربوية معينة كالمنهج الدراسي لابد وأن يتعرض إلى إعادة نظر موضوعية ومنظمة بين فترة وأخرى، فالمنهج الذي يمكن اعتباره مجموعة المواد المستخدمة من قبل الطلبة والمعلمين في العملية التربوية لابد وأن يخضع لعمليات تعديل وتنقيح من البداية وحتى مرحلة التطبيق العملي وخلال استمراره أيضاً للوقوف على الجوانب الإيجابية أوالسلبية فيه لأجل تحديدها ومعالجتها، وهنا يأتي دور التقويم وضرورته في هذه العملية من خلال إطلاع المقوّم على استخدام المواد المنهجية وملاحظة سلوك المعلم والمتعلم والإطلاع على المشكلات القائمة وكيفية علاجها، إذ قد يكتشف أن المواد صعبة على الطلاب، وأن الأنشطة أو الخبرات التعليمية غير مناسبة ولذلك يتم وضع خطة متكاملة لكيفية معالجة نقاط الضعف وتنظيمها في المنهج، أو إعادة تأليف الكتاب المستخدم بما يتناسب مع نتائج التقويم التي توضع لضمان سلامة المنهج وتعلم الطلبة بالشكل السليم. ومن هنا يرى الباحثون أن المناهج في أي مرحلة دراسية والكتب المؤلفة تخضع للتقويم والتغيير كل خمس أو عشر سنوات من تطبيقها لتماشي الواقع المعاش والتغيرات الاجتماعية والاقتصادية والتكنولوجية أو التطورات العلمية في

المجالات المختلفة، ولا يمكن لأي قطر يُريد من التربية أن تحقق أهدافه الاجتماعية أن يقف دون إحداث التغيرات الإيجابية المبنية على تقويم شامل وسليم لمناهجه الدراسية.

12) التوجيه المهني:

يستطيع المتخصص في التوجيه المهني استخدام اختبارات الاستعدادات والميول المهنية أو يجري مقابلات فردية أو جماعية مع الطلبة لفرض توفير المعلومات اللازمة عن كل طالب يمكن أن تساعد في توجيه ذلك الطالب نحو الدراسة التي تتناسب مع ميوله وقدراته، أو توجيهه الى المهنة التي تناسب رغباته لضمان نجاحه في تلك المهنة.

أنواع التقويم Types of Evaluation

هناك تصنيفات عدة للتقويم تختلف باختلاف الأغراض التي يستخدم فيها أو تبعاً لشمولية أو لأساليب تطبيقه أو توقيتاته وإلى غير ذلك. ونعرض في الصفحات التالية التصنيفات الشائعة للتقويم في الميدان التربوي.

أولاً: تصنيف التقويم وفقاً لأغراضه إلى:

أ. التقويم القبلي Pre Evaluation

ويقصد بهذا النوع من التقويم التأكد من الوضع الراهن للطلبة قبل البدء بعملية التدريس، والبدء بتخطيط هذه العملية وفقاً للخبرات التي يملكها الطلبة وبما يتناسب مع حاجاتهم. كما يتجسد هذا النوع من التقويم عندما نقوم بتوزيع الطلبة وفقاً لمستويات معينة في مجموعات تكون أكثر تجانساً، وأيضاً عندما يتم توزيع الطلبة على كليات وأقسام دراسية معينة وفقاً لاستعداداتهم وقدراتهم الخاصة. أي أن هذا النوع من التقويم يؤدي وظيفتين أحداهما (الكشف عن القدرات والاستعدادات Readiness) والأخرى هي (التعيين Placement).

ب. التقويم التشخيصي **Diagnostic Evaluation**

ويهدف إلى التعرف على نواحي القوة والضعف أو الخبرات السابقة عند المتعلمين من خلال إجاباتهم على اختبارات تشخيصية معينة تفسر نتائجها في ضوء وصولهم إلى مستوى معين يطلق عليه في العادة (مستوى الإتقان Mas-tery level)، حيث يفيد هذا النوع المعلم في تخطيط أساليب التعلم مع كل مجموعة أو طالب معين وفقاً لمستواه وقدراته التي تم تشخيصها مسبقاً.

ثانياً: تصنيف التقويم وفقاً لنوع الأساليب المستخدمة والبيانات والمعلومات

أ. التقويم الكمي **Quantitative Evaluation**

ويعتمد هذا النوع على النتائج الكمية (الرقمية) لأدوات القياس المستخدمة في التقويم كالاختبارات والاستبيانات وقوائم التقدير. كما أن هذا النوع يستخدم الأساليب الإحصائية في تحليل المعلومات الناتجة التي ينتج عنها إصدار احكام على البرنامج أو النشاط التربوي المقوّم.

ومن أشهر أنصار هذا الأسلوب كل من ستانلي Stanly وكامبل Camp-bell وآيكن Aieken وثورندايك Thorndike وغيرهم. ويتحمس أصحاب هذا الاتجاه، بل يتطرفون في استخدامه، حيث يؤكد كل من Campbell & Stanly على (أن الطرق الكمية التجريبية هي الطرق الوحيدة لترسيخ تقاليد علمية متراكمة تفيد في إجراء التحسينات في الموضوع المدروس من خلال التوصل إلى حقائق علمية دقيقة) بينما يتمادى البعض إلى أبعد من ذلك بقولهم: أن التجارب لا تؤدي إلى استنتاجات سببية فحسب بل تساعد في توضيح طبيعة المشكلة الاجتماعية قيد البحث، وأن التجريب هو طريقة الاختبار اللازم للحصول على معلومات صحيحة يُعوّل عليها ويمكن في ضوئها تخطيط البرامج الاجتماعية.

ب . التقويم النوعي **Qualitative Evaluation**

ويعتمد هذا النوع من التقويم على الملاحظات والآراء والانطباعات الشخصية والخبرة والمعرفة المتعلقة بالبرنامج أو النشاط الخاضع للتقويم مما قد يكون له فائدة في إكمال صورة البرنامج أو الظاهرة تحت التقويم، والتوصل إلى استنتاجات هامة حوله وتحديد قيمته، وقلما يعتمد هذا النوع على الجوانب الكمية والعمليات والأساليب الإحصائية المعقدة. ومن أنصار هذا الاتجاه كل من وايز ورين Weiss & Rain وبارلت وهملتون Parltt & Hamilton وكوبا Guba.

ويطرح أصحاب هذا الأسلوب استراتيجيات بديلة عن الأساليب الكمية يعتقدون بأنها أفضل من تصاميم التجارب كوسائل لتقويم البرامج ذات الأهداف الواسعة.

إن الطرق النوعية في التقويم تعتمد على أساليب معروفة منها على سبيل المثال: علم الأوصاف البشرية Ethnography ودراسة الحالة Case Study والمقابلات الشاملة Comprehensive Interview والملاحظات بالمشاركة Participant Observation وغير ذلك. ومن الجدير ذكره في مجال الفروق بين الاتجاهين الكمي والنوعي أن الجدل لا يزال حاداً بين أنصارهما حيث يؤكد الكميون "أن الأسلوب الكمي يشكل نظرة عالمية تتصف بالعملية والموضوعية، وله صفة استنتاجية Deductive في تعميم النتائج، بينما يرى النوعيون أن نظرتهم وأساليبهم تنسجم مع النظرة المتصفة بالاستقرائية Induc-tive والكلية Holic، والذاتية، والانثروبولوجية الاجتماعية، ويورد كل من كوك وريتشاردت (Cook & Reichardt, 1979) مقارنة لخصائص كل من الأسلوبين الكمي والنوعي في التقويم على النحو الآتي:

Quantitative Methods الطرق الكمية	Qualitative Methods الطرق النوعية
تبحث في الخصائص والأساليب المتعلقة بالظواهر الاجتماعية مع الأخذ بنظر الاعتبار بعض الحالات الذاتية للأفراد.	تهتم بفهم السلوك الإنساني من خلال الإطار العام الذاتي للفرد (الخبرة الذاتية).
تعتمد على قياسٍ محكم ودقيق (مُقيّد).	تستخدم ملاحظات طبيعية وغير مقيدة.
موضوعية في أساليبها.	ذاتية في أساليبها.
تعتمد على الأسلوب الاستنتاجي، الاستدلالي.	ذات طابع اكتشافي وصفي، استقرائي.
أساليبها موجهة نحو النتائج. نتائجها قابلة للتعميم لأنها تعتمد على	أساليبها موجهة نحو العملية أي (الأسلوب أو الطريقة).
نتائجها قابلة للتعميم لأنها تعتمد على حالات متعددة.	نتائجها غير قابلة للتعميم في أغلب الأحيان لأنها غالباً ما تكون دراسة حالات محددة.
تستخدم القياس المحكم المحدد في ملاحظاتها.	تعتمد على ملاحظات طبيعية غير مضبوطة منهجياً
أساليبها تهتم بالدقائق والتفاصيل المتعلقة بالظاهرة أو المشكلة.	أساليبها تهتم بالأشياء الكلية للظاهرة أو المشكلة.

ولابد من الإشارة هنا إلى أنه ورغم أن لكلا الطريقتين الكمية والنوعية فوائد في عملية التقويم إلا إنه لا يمكن الاستغناء عن أي منهما بل وعلى العكس من ذلك فإن بالإمكان المزج بين محاسن الأسلوبين عندما تتطلب عملية التقويم وظروفها ذلك. فالمقوّم مهما كان مؤمناً بالأساليب الكمية وتطبيقاتها لابد وأن يعتمد في تفسيره للنتائج أو تبريرها على الأساليب النوعية.

كما أن هناك حالات عديدة في مجالي البحث والتقويم يصعب فيها تطبيق الأساليب الكمية والتجريبية تطبيقاً كلياً وحيداً دون اللجوء إلى الأساليب النوعية للتوصل إلى استنتاجات دقيقة ومقبولة، ولذلك يمكن القول أن أفضل أساليب التقويم تلك التي تجمع بين الأساليب الكمية والنوعية في نفس الوقت وفقاً للظروف التي تجري فيها عملية التقويم وطبيعة الموضوع المقوّم.

ولتوضيح كيفية استخدام الأساليب الكمية والنوعية معاً في عملية التقويم يمكننا أن نسوق المثال الآتي:

لنفرض أن أحد المقومين الكميين يريد دراسة موضوع عزوف بعض الطلبة عن المشاركة في الأنشطة الرياضية فإنه لابد أن يهتم بجملة من الأمور المتعلقة بتحديد مشكلة البحث وأهدافه والمجتمع الذي يريد أن يعمم إليه نتائج البحث وحجم عينة البحث والأساليب الإحصائية المطلوب استخدامها والأدوات المستخدمة وإيجاد صدقها وثباتها إلى غير ذلك. وفي نهاية بحثه سيتوصل الباحث إلى نتائج رقمية تحدد نسب الموافقين والمعارضين، يدعمها بنتائج بحوث سابقة وعندما لا يجدها يلجأ إلى استخدام التفسيرات المنطقية التي تتفق مع المعايير الاجتماعية والعلمية، أو تلك الناشئة عن الخبرة والملاحظات والانطباعات الشخصية وهنا نرى الباحث زاوج بين الأسلوبين الكمي والنوعي في نفس الدراسة.

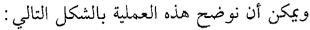
أما المقوم أو الباحث النوعي الذي يريد القيام بدراسة نفس الموضوع فإنه قد يلجأ إلى أساليب أقل تعقيداً من تلك التي استخدمها الباحث الكمي حيث يلجأ إلى الملاحظات الطبيعية ويرى هل أن أكثرية الطلبة يشارك في الأنشطة الرياضية ويجري مقابلات معمقة وشاملة مع بعض الطلبة ذكوراً وإناثاً في الجامعات المختارة كميدان لدراسته ويُثير من خلال تلك المقابلات مناقشات معمقة للتوصل إلى الآراء الشخصية حول الإقدام أو الإحجام عن المشاركة ليكوّن ملفاً واسعاً عن الموضوع ويربط المشكلة بعوامل اجتماعية واقتصادية مؤثرة آخذاً بنظر الاعتبار طبيعة المجتمع الذي تجري فيه الدراسة، مستعيناً بخبرته الشخصية وانطباعاته عن الموضوع فهو كان طالباً والآن أستاذاً في الجامعة بما يجعله ملماً بجوانب الموضوع المتعددة، وقد يستخدم في عرض نتائج دراسته إضافةً لما ورد ذكره بعض الأرقام والنسب المئوية التي تُبرز مشكلة البحث وإعداد الموافقين والرافضين بما يعني مزاوجته للأسلوبين الكمي والنوعي في بحثه.

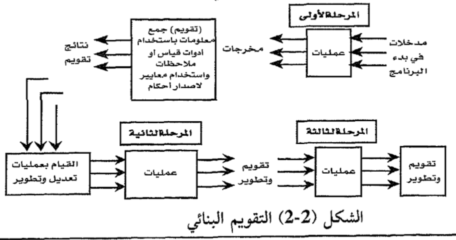
ثالثاً: تصنيف التقويم وفقاً لإجراءات التوقيت إلى:

أ. التقويم البنائي (التكويني) Formative Evaluation

يلاحظ من عنوان هذا النوع أنه يستهدف البناء أو التطوير حيث يجري ويستمر مع استمرار عمل البرنامج التربوي أو العملية التعليمية ويُنفَذ بشكل دوري بعد أن تقسم الفترة الخاصة بالبرنامج إلى عمليات وتقويم وتغذية راجعة. ففي عملية التدريس يمكن للمعلم أن يقسم عمله إلى فترات، وبعد كل فترة تدريس يقوم بعملية تقويم أي مراجعة تدريسه للفترة المنصرمة من حيث طريقة التدريس وفاعليتها وتحديد نواحي القوة والضعف لدى طلبته مما يلقي على هذا النوع طابعاً تشخيصياً تعقبه عملية تصحيح تهدف إلى وضع استراتيجيات وأساليب لمعالجة نواحي الضعف وترصين نواحي القوة، بما يؤدي الى تطوير العملية في المرحلة اللاحقة وهو ما يسمى التغذية الراجعة Feedback. وفي عملية تقويم البرامج والمشاريع التربوية يمكن توضيح هذا النوع بالمثال الآتي:

لنفرض أننا بصدد تقويم برنامج تربوي معين وهنا لابد من تقسيم فترة عمل البرنامج إلى ثلاث فترات زمنية محددة يجري في نهاية الفترة الأولى تشخيص نواحي القوة والضعف باستخدام أدوات تقويمية متنوعة ثم يتم وضع أساليب لمعالجة نواحي القصور وترصين نواحي القوة بناء على نتائج التقويم، واستخدام تلك المعالجات في الفترة اللاحقة، مما يضفي على الفترة اللاحقة إمكانية التطوير الناشئ عن عملية تقويم المرحلة السابقة، وكأن المرحلة اللاحقة يتم تشكيلها أو بناءها على أساس الاستفادة من نتائج تقويم المرحلة السابقة.


ويمكن أن نوضح هذه العملية بالشكل التالي:



الشكل (2-2) التقويم البنائي

ومن خلال ما سبق يمكن أن نستنتج أن التقويم البنائي يتميز بما يأتي:

1- أنه عملية مستمرة طوال استمرار البرنامج المقوّم.

2- يحتاج إلى جهود تتمثل بإعداد أدوات قياس واختبارات ووضع معايير ومحكات تؤدي إلى صياغة أحكام عامة عن البرنامج.

3- أن هذا الأسلوب يؤدي إلى اتخاذ قرارات تطويرية وليس قرارات عقابية.

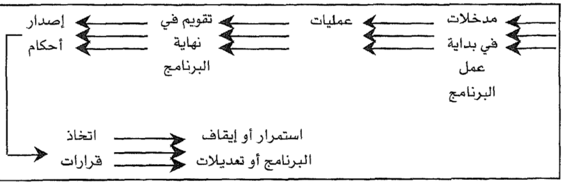
4- أنه يحتاج إلى وقت أطول، وأشخاص لديهم خبرات ميدانية وعملية في تطوير أدوات البحث والقياس وتحليل البيانات وإمكانات خاصة لعمل أحكام موضوعية ودقيقة.

ب. التقويم الختامي (النهائي) Summative Evaluation

ويقصد به ذلك التقويم الذي يُستخدم في نهاية تطبيق برنامج معين، حيث يجمع المقوّم بيانات ومعلومات تتعلق بسير عمل البرنامج منذ بداية تنفيذه وحتى فترة معينة، ويجري تحليل تلك البيانات ومقارنتها بمعايير محددة، ثم يصدر أحكاماً تهدف إلى التأكد من مدى تحقق أهداف ذلك البرنامج لأجل اتخاذ قرارات بشأن استمراره أو إيقافه. وقد ينشأ عن هذا النوع من التقويم أيضاً اتخاذ قرارات إدارية تتضمن العقوبة أو النقل للأفراد العاملين في ذلك البرنامج.

وفي العملية التدريسية يتضح هذا النوع من خلال تطبيقه مع نهاية فصل دراسي أو سنة دراسية للتأكد من مدى تحقق الأهداف التي وضعت في بداية العمل. كما يمكن أن يتضح تطبيق هذا النوع من خلال الحكم على مسيرة الطالب طوال سنة دراسية أو فصل دراسي بعد أن يجري تقويم تحصيله باختبارات تحصيلية نهائية في أواخر العام أو الفصل الدراسي كما هو معمول به في بعض الجامعات العالمية عندما يقدم الطالب امتحانات نهائية فقط يتم بعدها الحكم على مستواه واتخاذ القرار بشأن نقله إلى صفٍ أعلى أو بقاءه في صفه فترةٍ أخرى.

والمخطط الآتي يوضح عملية (التقويم الختامي):



الشكل (3-2) التقويم الختامي

ويشير بعض الباحثين ومنهم (سعادة وإبراهيم، 1997) أن هذا النوع من التقويم يُستخدم في نهاية فترة زمنية أو نهاية مادة دراسية أو برنامج دراسي لأغراض منح الشهادات أو تقويم مدى التقدم الذي تمّ تحقيقه في التعليم من حيث فاعلية منهج أو مادة دراسية أو خطة تربوية محددة.

رابعاً: تصنيف التقويم حسب شموليته إلى:

أ. التقويم المكبر (الكلي) (الشامل) Macro Evaluation

ويركز هذا النوع على مخرجات النظام. أي تقويم الحالة وفق منظورها أو إطارها العام، وليس وفق حالتها الراهنة وبكلمة أوضح تقويم الفعالية التربوية الجزئية وفقاً لعلاقتها بالحالة العامة. فعندما نقوّم نشاطاً تربوياً أو مشروعاً تربوياً يُفترض أن نقوّم ذلك النشاط بمدى ما يقدمه ويساهم به في تحقيق الأهداف التربوية العامة في قطر معين. كما أن هذا الأسلوب يأخذ بعين الاعتبار جهود العاملين في العملية التربوية (عناصرها الهامة) كالمعلمين والمشرفين والمدراء والمرشدين والطلبة وأولياء الأمور إضافةً إلى المواد التعليمية المساهمة في إنجاز العملية التربوية لما لهذه العناصر من دور كبير وفاعل في تحقيق الأهداف العليا للتربية المتمثلة في تحقيق النمو المتكامل للفرد والمجتمع.

إن هذا النوع من التقويم يتطلب جهداً كبيراً وقد يستغرق فترة زمنية طويلة وإمكانات مادية كافية حيث يكون العمل فيه بشكل فريق، لذلك يكثر استخدامه في مؤسسات التخطيط القومي التي تعتمد على استراتيجيات التطوير البعيد المدى.

ب. التقويم المُصغّر (الجزئي) Micro Evaluation

ويسمى أيضاً (التقويم الجزئي) حيث يتناول جانباً محدداً من جوانب العملية التربوية مثل تقويم المعلم لتحصيل طلبته، وتقويم كفاءة المعلم في التدريس وتقويم طريقة تدريس معينة أو تقويم نوعية الأسئلة الامتحانية التي يضعها لطلبته أو تقويم مكتبة المدرسة وغير ذلك كلها أمثلة لهذا النوع من التقويم.

وينتقد بعض المقومين والباحثين هذا النوع بأنه نسبي في شموليته إذ يعتمد على ظروف خاصة محددة قد لا تنطبق على حالات مشابهة أخرى. كما أنه لا يوفر معلومات كافية، ويصعب أحياناً تفسير نتائجه تفسيراً دقيقاً لذلك تعزى النتائج إلى احتمالات أو ظروف معينة للموقف المقوّم مما يجعل التقويم غير واضح ويصعب تعميم نتائجه إلى حالات مشابهة. إضافة إلى صعوبة تطبيق مبدأ (المسائلة Accountability) أي تحديد الجهة المسؤولة عن عدم تحقيق الأهداف.

خامساً: تصنيف التقويم حسب الجهة التي تقوم به إلى:

أ. التقويم الخارجي: External Evaluation

وهو التقويم الذي تقوم به جهات أو هيئات أو مؤسسات تقويمية مستقلة ولا علاقة لها بالمؤسسة أو البرنامج المطلوب تقويمه في غالب الأحيان. وربما تساعد هذه الاستقلالية في توفير الحياد والموضوعية في عملية التقويم. وكثيراً ما تطلب المؤسسات التعليمية ولا سيما الجامعات من الجهات المتخصصة بإجراء عملية تقويم برامجها من أجل الحصول على شهادة محايدة أو تحقيق ما يسمى بالاعتماد أو (إجازة الاعتماد) التي تؤكد السمعة العلمية الجيدة لتلك الجامعة لتأخذ مكانها بين الجامعات العالمية المرموقة. كما أن التقويم الخارجي يمكن أن يقدم صورة شاملة لصانع القرار لاتخاذ قرارات تتعلق بالسياسات والتخطيط وبيان أثر البرنامج والمشروعات التربوية، وإنتقاء الأفراد وتحقيق مبدأ المساءلة.

ب. التقويم الذاتي أو الداخلي Self Evaluation

ويقصد به قيام المؤسسة أو المدرسة بمبادرة ذاتية لإجراء عملية التقويم

الداخلي لأنشطتها في المجالات المختلفة سواء المتعلقة بالمعلم أو المتعلم أو الإدارة بهدف إصدار أحكام موضوعية عن أداء تلك المؤسسة. ولا يقتصر التقويم الذاتي على المؤسسات فحسب بل يمكن للأفراد داخل تلك المؤسسات من إجراء التقويم لأنفسهم، مثال ذلك: التقويم الذاتي للمعلم من حيث طريقة تدريسه، شخصيته، معرفته للأهداف التي يسعى إلى تحقيقها، كفاءته العلمية وتعامله مع الطلبة والأساليب التي يتبعها لتحفيزهم نحو التعلم وإلى غير ذلك.

ونظراً لأن التقويم الذاني بدأ بالانتشار في المؤسسات التربوية في العالم المتقدم وفي بعض الأقطار العربية ومنها سلطنة عُمان حيث بدأت وزارة التربية والتعليم بتطبيقه تجريبياً على بعض المدارس الحكومية ابتداءً من العام الدراسي 2003 - 2004، فإن المؤلفين يَرونَ أن من الضروري التوسع في تناول هذا النوع من التقويم بشكل مفصل.

ومن الجدير بالذكر أن التقويم الذاتي Self Evaluation لا يختلف كلياً من حيث أسسه ومنهجيته وأهدافه عن التقويم الخارجي External Evaluation الذي تقوم به جهة من خارج المؤسسة. إلا أن العاملين في ميدان التقويم يضعون جملة فروق وأفكار في نواح أخرى منها:

1- إن فلسفة التقويم الذاتي تنطلق من مسلمات أهمها أن المؤسسة أعرف من غيرها بما يدور في داخلها، وما هي طبيعة أهدافها، وما هي الوسائل التي تتبع لتحقيق تلك الأهداف، كما أن لديها القدرة الأكبر على تشخيص نقاط الضعف داخلها وكيفية وضع المعالجات لتصحيحها، وما هي نقاط القوة وكيف تتم المحافظة عليها واستمرار ديمومتها.
2- هناك من يطرح آراء أخرى يشكك فيها بموضوعية التقويم الذاتي من حيث أن المقوم فيه يحاول غض النظر عن نقاط الضعف والأخطاء التي تتواجد في تلك المؤسسة، وإظهارها زوراً أنها تتمتع بمقومات النجاح والقوة في جميع جوانبها وعناصرها. وهذا يستدعي عدم الاعتماد على عملية التقويم الذاتي، مما يُبرر ضرورة استخدام التقويم الخارجي.

3- يطرح البعض الآخر رأياً يستند إلى أن الإنسان أو اية منظمة أو مؤسسة لا يمكن لكل منهم أن يرى عمله ونشاطه بشكل واضح إلا من خلال الآخرين، ومن أجل منع التحيز وضبابية الحقائق لابد من التقويم الخارجي لتحديد نقاط الضعف وجوانب القوة، ومن ثم إصدار الحكم على نشاطات تلك المؤسسات بناءً على معايير محددة.

4- يُضاف إلى ما ورد أعلاه، هناك من يطرح رأياً آخر يفيد بأن التقويم الخارجي أكثر دقة وموضوعية من التقويم الذاتي، لأنه يجري من قبل خبراء متخصصين في عملية التقويم، بينما يعتمد التقويم الذاتي (لا سيما في المدارس) على أفراد أقل معرفة وتخصصاً في عملية التقويم على الرغم من قيام بعض المؤسسات بتدريبهم بشكل سريع حول كيفية إجراء التقويم لكنهم قد لا يصلوا إلا إلى معرفة أو قدر من الموضوعية، سواء في تخطيط أو تنفيذ عملية التقويم التي قد تصبح من خلالهم عملية آلية غير دقيقة في كثير من الأحيان.

5- يذكر ستفلبيم (1981) أن الظروف المحيطة بالبرنامج أو بعملية التقويم هي التي تحدد نوع الاستراتيجيات الأفضل لاستخدامها في عملية التقويم، هل تستخدم التقويم الذاتي (الداخلي) أم التقويم الخارجي، ويوصي آخرون ومنهم Suchman أن الجمع بين التقويم الداخلي والخارجي هو الأسلوب المفضل للوصول إلى نتائج تقويمية دقيقة وموضوعية وخططاً تطويرية هامة.

يتضح مما تقدم أن التقويم الداخلي والتقويم الخارجي يكملان بعضهما ويوصلان إلى نتائج إيجابية، فالجوانب التي لم تغطى بالتقويم الداخلي (معروفة أم متخطاة) يمكن أن يكشفها ويغطيها التقويم الخارجي، أما النواحي التي قد لا يستطيع التقويم الخارجي أن يدركها فإن عملية التقويم الذاتي هي التي توضحها أمام المقوّم الخارجي من خلال تقرير التقويم الذاتي الذي يقدم للمقوم الخارجي فيزيد معرفته ويزيل الغموض عن تلك الجوانب. فإذا كان مدير المدرسة أو عميد الكلية هو الشخص المسؤول عن سير التقويم الذاتي،

فإن المقوم الخارجي يفترض أن يلتقيه قبل بدء التقويم ليناقشا سوية مسألة الأهداف، وخطة التقويم والاستراتيجيات التي سيتم اتباعها، وحتى التطرق إلى أدوات التقويم التي ستستخدم في جمع البيانات عن المدرسة أو الكلية، وسوف يكمل كل منهما أي نقص أو غموض في المعلومات الأخرى مما يؤدي إلى نجاح العملية وتقدمها. كما أنهما يجب أن يتفقا على المعايير الممكن استخدامها لإصدار الأحكام، فقد يحدث أن المقوّم الخارجي يهتم بالأهداف الكلية بينما يكون اهتمام المدير بالأهداف الآنية والتركيز على الجهود المبذولة والأساليب المستخدمة في العمل. إضافة لذلك يمكن للطرفين أن يتفقا على أهمية التوصل إلى نتائج تخدم في إطارها النهائي عمل المؤسسة وتطويره، وأن يتم التركيز في كتابة التوصيات النهائية على وضوحها وسهولة تطبيقها، ووضوح الآليات الخاصة بتنفيذها بحيث تكون متوافقة مع قدرات المؤسسة المالية والمعنوية لكي يتم تنفيذها بسهولة بما يؤدي إلى التطوير الهادف.

فلسفة ومبادئ التقويم الذاتي:

يعتمد التقويم الذاتي على مجموعة من المبادئ العامة للتقويم ومنها:

1- لابد للتطوير أن يبدأ من الذات، فذلك أفضل من التطوير الذي يأتي من الجهات العليا، لأن التطوير الآتي من الذات يوفر قناعة أكبر واندفاعاً أكثر نحو تحقيق أهداف المؤسسة.

2- لابد لكل مؤسسة تربوية أو غير تربوية تحرص على تطوير أنشطتها أن تمتلك الجرأة في كشف نواحي الضعف في العمل وتحديدها بوضوح من أجل دراسة أسبابها ووضع المعالجات الخاصة بتطويرها قبل أن تسلط الضوء على نواحي القوة في برامجها.

3- أن التقويم الذاتي، ومن خلال المعرفة الدقيقة للقائمين عليه وهم جزء من المؤسسة المقوّمة يستطيعون أن يركزوا بشكل أكثر دقة من غيرهم على الجوانب التي تحتاج إلى تقويم في العمل المدرسي، واختيار الأساليب والأدوات التقويمية التي تتناسب وظروف تلك المؤسسة من أجل توفير معلومات وبيانات مفيدة في عمليتي الحكم والتطوير الذي يعقب إصدار الحكم.

4- لابد لعملية التقويم الذاتي أن تشمل جميع جوانب وعناصر العملية التربوية كالتعليم والتعلم، والمنهج، والمستلزمات المدرسية المصاحبة لها، والطالب والمعلم وقيادة أو إدارة المدرسة، حتى تتصف بمبدأ الشمولية الذي يعتبر واحداً من الخصائص المهمة لعملية التقويم.

5- إذا ما أريد لعملية التقويم الذاتي أن تحقق أهدافها بشكل دقيق يجب أن تتصف بالموضوعية في جميع خطواتها ابتداء من التخطيط لها وحتى كتابة التقرير النهائي، والابتعاد الكامل عن تحاشي التركيز على نقاط الضعف والاهتمام بإبراز جوانب القوة، لأن ذلك يعني الإخلال بأخلاقيات البحث والتقويم العلمي وسوف لن يوصل المؤسسة قيد التقويم لتحقيق الغايات المتوخاة من التقويم وهي التطوير والتحسين وضمان جودة الأداء.

أهداف التقويم الذاتي:

على الرغم من أن أهداف التقويم الذاتي لا تبتعد كثيراً عن أهداف التقويم بجميع أنواعه لكنها تتميز عنها في الآتي:

1- تشجيع المؤسسة أو المدرسة للتدرب على تطوير نظام تقويمي ذاتي فاعل خاص بها ويتفق مع ظروفها.

2- يوفر الفرصة للعاملين في المؤسسة بالمشاركة الفاعلة واكتساب الخبرة باساليب التقويم وإجراءاته، وما ينشأ عن ذلك من المشاركة والتحمس في تحسين ورفع كفاءة الأداء في المدرسة التي يعملون فيها.

3- يشجع العاملين في المؤسسة على دعم عمليات التطوير والتغيير التي جاءت كنتيجة لعملية التقويم التي شاركوا فيها أنفسهم بما ينعكس إيجابياً على تحسين كفاءة العمل والنشاطات المدرسية في جميع مجالاتها.

4- إن التقويم الذاتي الذي يؤدي إلى تغيرات إيجابية في العمل المؤسسي أو المدرسي يُحسن من صورة المدرسة أمام المجتمع المحلي ويزيد من ثقة أولياء الأمور بإدارة المدرسة ومعلميها، لا سيما بعد أن يلمسوا نتائج التطوير، ويعلموا أنها جاءت بجهود العاملين في المدرسة ومشاركتهم في عملية التقويم الذاتي.

وتؤكد بعض الأدبيات أن التقويم الذاتي يقع في قلب الإدارة الجيدة والتطوير المدرسي الفاعل، وأن هذا التقويم يُبنى على مؤشرات عن المستويات التي يحققها الطلاب ونوعية التعليم (التدريس) والإدارة والقيادة الفاعلة. كما أن استخدام الشواهد المتحققة في هذه المجالات الثلاثة تمكن المسؤولين القياديين في المدرسة ليكونوا أكثر دقة في تحديد الأولويات والأهداف المتعلقة بالتطوير. كذلك تعتبر هذه الشواهد عوامل أساسية لتوفير نقطة الانطلاق للتقدم الممكن قياسه سواء كان بشكل تحسن سابق، أو عندما نريد مقارنة أداء المدرسة بأداء مدارس أخرى مشابهة لها، كما أن رفع مستوى تحقيق التميز والمحافظة عليه يتطلب تخطيطاً مدرسياً يستهدف إجراء تغييرات في الإدارة ، والتدريس الذي يمكن الطلاب من بذل جهودهم بشكل أكبر، ويمكن تفعيل ذلك من خلال قيام المسؤولين عن قيادة المدرسة من وضع استراتيجيات التحسين الواضحة الخطوات لتصبح جزءاً من دورة التخطيط الاستراتيجي المستمر.

خصائص التقويم الذاتي الفعال:

يؤكد بعض العاملين في ميدان التقويم التربوي على جملة خصائص يجب أن يتميز بها التقويم الذاتي للمؤسسات التربوية ومنها:

1- أن يكون صادقاً أي يعكس الواقع الفعلي بشكل واضح دون التغاضي عن نواحي القصور وإبراز نواحي القوة فقط.

2- أن يهتم بعملية التخطيط الاستراتيجي الخاص بالتطوير والتحسين اللاحق ويضمن جودة الأداء.

3- يُشرك جميع العاملين في الملاك المدرسي وبجميع مستوياته في تحديد بيانات كافية عن المخرجات والأداءات بشكل دقيق وموضوعي.

4- يهتم بشكل خاص بآراء الطلاب والمستفيدين من المدرسة، العاملين والآباء والمجتمع المحلي ويأخذها بنظر الاعتبار في عملية التقويم.

5- يهتم بإصدار أحكام موضوعية وواقعية في ضوء الأداء المقاس، وأن تكون متحررة من التحيز وشاملة تعتمد على ملاحظات مباشرة للواقع بكل جوانبه ومستندة إلى محكات أساسية واضحة.

6- يقود أو يؤدي إلى خطوات تطوير موضوعية واقعية تسترشد بأهداف واضحة ومعايير محددة.

التخطيط للتقويم الذاتي:

تمثل المدرسة مجموعة منظمات مركبة متعددة، ولأجل التخطيط اللازم لتحسين أدائها فعليها أن تأخذ بالاعتبار مجموعة متنوعة من المخرجات منها:

(نتائج الاختبارات، تقييمات المعلم، سلوك الطالب واتجاهاته، أعمال الطلاب) وعلى المقومين أن ينظروا إلى المدرسة من وجهات نظر مختلفة تتمثل في: (الآباء، الطلاب، المعلمين، ومجموعة من الطلاب الذين يشعرون أنهم مبعدين من نشاطات الحياة المدرسية، وأخيراً أشخاص من خارج المدرسة من الذين لديهم تصورات عن بعض مجالات العمل المدرسي مثل العمال، المستشارين والكتبة).

كما أن على المدرسة أن تستخدم ما تعرفه عن ما يحدث في الصفوف لتشخيص جوانب القوة والمجالات الخاصة بالتطوير، وعلى وجه الخصوص يجب أن تأخذ بالاعتبار أن كل وجه أو جزء من نشاطات المدرسة يؤثر على جميع الطلاب أو مجموعة منهم.

ومن أجل أن نضمن نجاح عملية التقويم الذاتي علينا نعرف أو نحدد المجالات التي يجب أن تشملها عملية التقويم ونركز على نتائج التعليم، كما أن إصدارالأحكام على أساس الأدلة المتوفرة سيساعد المقوم على تشخيص نواحي القوة والضعف في البرنامج وتحديد المجالات التي تحتاج إلى تحسين، ولذلك يجب أن يركز المقوّم على ما يأتي:

أولاً: المخرجات:

1- مدى تحقيق معايير محددة.

2- أعداد المسجلين في المدرسة.

3- التقدم الحاصل عند الطلاب، وهنا نحتاج إلى قياس ما إذا كانت

مجموعات الطلاب المختلفة قد حققت قدراً كافياً من التعلم الذي يستهدفون تحقيقه.

4- تحصيل الطلاب في مجالات أخرى من الحياة المدرسية مثل: الخبرة العملية.

5- التحصيل في الجوانب الرياضية والفنية.

6- سلوك الطلاب: اتجاهاتهم، قيمهم والتطور الشخصي لهم.

ثانياً: الاستعدادات:

تقويم نوعية الأوجه المختلفة للاستعدادات المدرسية المقدمة للطلاب حيث أنها تساعد المقوّم على تحديد ما إذا كانت هناك حاجة للتحسين وما إذا كانت هذه الحاجة ملحة لغرض إنجازها، لذلك فإن التقويم يجب أن يركز على:

1- نوعية التدريس في عموم المدرسة.

2- نوعية الإدارة والقيادة التي يظهرها المسؤولون في المدرسة.

3- نوعية المناهج.

4- نوعية العلاقة بين الإدارة المدرسية وأولياء الأمور والمجتمع المحلي.

5- المناخ التعليمي والبيئة المدرسية.

العناصر التي يشملها التقويم الذاتي:

لا تختلف العناصر التي يشملها التقويم الذاتي عن تلك التي يشملها التقويم بشكل عام خارجياً كان أو ذاتياً. فالعناصر المادية والبشرية يشملها التقويم، المعلم، ومدير المدرسة، والمشرف التربوي، والموجه، والطالب كلها عناصر تشارك مشاركة فاعلة في العملية التربوية لذلك فإن نجاحها أو فشلها يمثل مستوى أداء تلك العناصر لذلك علينا أن نقوّم هؤلاء تقويماً متمثلاً بأعمالهم التي يقومون بها لكي نحكم على نجاح أو فشل تلك الأعمال.. أما العناصر المادية المساهمة في العملية التربوية فهي كثيرة ومتنوعة وتعتبر وسائل هامة تساهم في تحقيق أهداف العملية التربوية أو فشلها في كثير من الأحيان،

وتُسرع في تحقيق أهداف العملية أو تباطؤها، فالمنهج المدرسي والكتاب المدرسي ونوعه وإخراجه وتناسب مستواه مع المستوى العقلي للطالب، ومصادر التعلم بكافة أنواعها، وحتى البناية المدرسية ومحتوياتها وأقسامها تشملها عملية التقويم والتعديل اللاحق، ولا يمكن أن يعزل أي من العناصر المادية والبشرية بعيداً عن عملية التقويم، لكن عملية التقويم الذاتي في بعض الأحيان لا تمتد إلى جميع العناصر الداخلة في العملية التربوية، بل تحدد أهم تلك العناصر وأداءها، كما نلاحظ في إجراءات التقويم الذاتي المطبقة في مدارس السلطنة حيث تم تحديد ثلاثة عناصر مهمة تخضع للتقويم الذاتي وهي عملية التعلم والتعليم والإدارة المدرسية، بعناصرها البشرية المتمثلة بالطالب، المعلم، ومدير المدرسة، ولم تقف العملية عند هذا الحد بل أمتدت إلى إشراك عنصر آخر مهم هو أولياء أمور الطلاب الذين يمثلون عناصر هامة تعكس وجهة نظر المجتمع المحلي في الأداء المدرسي.

مراحل وإجراءات التقويم الذاتي:

اهتمت العديد من الدول المتقدمة بتأسيس أنظمة التقويم الذاتي للمدرسة School Self Evaluation والتحقق من مصداقية هذا النظام استناداً إلى استخدام نوع آخر من التقويم هو التقويم الخارجي External Evaluation، حيث يكمل أحدهما الآخر للتوصل إلى نتائج إيجابية من أجل تطوير وتحسين عمل المدرسة، وهذا ما يجري الآن في (مشروع تقويم الأداء المدرسي) الذي تنفذه وزارة التربية والتعليم في سلطنة عمان.

وعلى الرغم من تعدد الأساليب ونماذج التقويم في الميدان التربوي إلا أنها لا تختلف عن بعضها البعض اختلافات جوهرية بل ربما يختار كل نموذج أسلوباً مختلفاً لتحقيق أهدافه وأدوات كمية أو نوعية أو مؤلفة من النوعين ولكل نموذج معايير خاصة يعتمدها للحكم على الوقائع من أجل التوصل إلى نتائج تؤدي إلى الكشف عن جوانب القوة والضعف، ووضع أساليب ومعالجات التطوير اللاحقة، وكمثال تطبيقي حي على عملية التقويم الذاتي المدرسي نقدم في الصفحات القادمة عرضاً ملخصاً للإجراءات والمراحل التي

يتم استخدامها في عملية تقويم الأداء في مدارس السلطنة حيث تتضمن ستة مراحل أساسية هي:

المرحلة الأولى: جمع البيانات:

تتطلب عملية جمع البيانات والمعلومات في جميع أنواع التقويم جانبين أساسيين الأول تحديد العناصر التي يشملها التقويم وفي مشروع تقويم الأداء المدرسي في السلطنة ثم تحديد ثلاثة عناصر هي (التعلم، التعليم، والإدارة المدرسية).

أما الجانب الثاني في هذه المرحلة فيتضمن تحديد الأدوات التي تستخدم لجمع المعلومات، وفي ذات المشروع يقوم المسؤولون عن التقويم باستخدام استمارات خاصة تم تصميمها من قبل خبراء متخصصين بإشراف وزارة التربية والتعليم وهي كالآتي:

1- **التعلم**: فيما يخص هذا العنصر يتم تقويمه من خلال معيارين هما (التحسن في أداء الطلاب، وجودة أعمال الطلاب). ويتم جمع المعلومات الخاصة بمؤشرات التحسن في أداء الطلاب من خلال: تحليل أعمال الطلاب، الملاحظات الصفية، الإطلاع على سجلات قياس المعلمين لأداء الطلاب، تحليل نتائج الامتحانات، مقابلة المعلم الأول والمناقشات التي تجري مع المعلمين حول أداء الطلاب.

أما الأدوات التي تستخدم للحصول على المعلومات من المصادر السابقة الذكر فهي:

أ. **استمارة الملاحظة الصفية**: التي يقوم بتنفيذها مدير المدرسة ومساعدهُ والمعلمون الأوائل والمعلمون أثناء تبادل الزيارات.

ب. **استمارة تحليل أعمال الطلاب**: ويقوم بتنفيذها مدير المدرسة ومساعده والمعلمون الأوائل وعلى مدى مرتين في كل فصل دراسي والهدف منها تكوين صورة عن عملية التعلم ومستوى أعمال الطلاب ويركز فيها على التحصيل والإنجاز والاتجاهات والسلوك.

2- التعليم (التدريس): ويتم تقويمه من خلال معيارين هما:

أ. تأثير نوعية التدريس على مستوى تعلم الطلاب، وهو مؤشر لجودة التدريس.

ب. تأثير التدريس في تلبية احتياجات التعلم الخاصة بجميع الطلاب.

وذلك من خلال المصادر الآتية:

- الملاحظة الصفية.
- تحليل أعمال الطلاب.
- الإطلاع على سجلات المعلمين.
- مقابلة المعلم الأول.
- التحدث مع المعلمين.
- التحدث مع الطلاب.

3- الإدارة المدرسية: يتم تقويم هذا العنصر من خلال المعايير الآتية:

- العمل على تحقيق مستويات تعلم عليا لدى الطلاب.
- الإشراف على عمليتي التعلم والتعليم.
- فاعلية المدرسة في رعاية الطلاب.
- تفعيل العلاقة مع أولياء الأمور والمجتمع المحلي.
- التوظيف الفاعل لكادر المدرسة والمبنى المدرسي.

ويتم تحقيق ذلك من خلال:

- استبانة أولياء الأمور التي تقوم إدارة المدرسة بتوزيعها على عينة من أولياء الأمور.
- استبانة آراء الطلاب، ويوزعها الأخصائي الاجتماعي على عينة من الطلاب.
- استمارة آراء المعلمين في الإدارة المدرسية، ويقوم مدير المدرسة أو مساعده بتوزيعها على عينة من المعلمين في نفس المدرسة.
- الإطلاع على وثائق المدرسة وسجلاتها وملفاتها.

- التوظيف الفعال لمصادر التعلم (مكتبات، مختبرات، مركز مصادر التعلم)، ويقوم المعلم الأول لكل مادة بتعبئة الجزء الخاص به في هذه الاستمارة وإعطاء أحكام وفق سلم تقدير للاستجابة هو (جيدة - مرضية - غير مرضية).
- دراسة خطط المدرسة.
- الأنشطة المدرسية (الخارجية والداخلية).
- متابعة المبنى المدرسي بجميع مرافقة (ملاعب، دورات مياه، برادات ماء، الجمعية التعاونية... إلخ) ومدى تفعيلها (وزارة التربية والتعليم، 2004).

المرحلة الثانية: تحليل البيانات

بعد الانتهاء من تجهيز البيانات وتنظيمها يتوجه الانتباه إلى عملية تحليل البيانات Data Analysis باستخدام أساليب مناسبة قد تستدعي استخدام اساليب إحصائية. وحيث أن هذه البيانات جهزت من خلال أدوات بحث بسيطة استخدمت في عملية التقويم الذاتي لا يشترط استخدام وسائل إحصائية معقدة معها كتلك التي تستخدم في البحوث العلمية المعدة للنشر، هذه الأساليب البسيطة ضرورية لاستخلاص معلومات مفيدة من البيانات تحقق أغراض التقويم من جهة، ويمكن فهمها بسهولة من قبل القارئ العادي وتوضح صورة مقبولة مبسطة لصناع القرار. والبيانات التي تجمع تعتبر بيانات خام لم تخضع للتقييم بعد، وعندما يجري تجهيزها وتحليلها تصبح في سياقات ذات معنى تمهد لعمل أحكام وفق معايير معينة. ولعل المهمة الأولى لفريق البحث الذي قام بجمع البيانات هي تحديد المتغيرات الرئيسية وتصنيفها من حيث نوعها ومستوى قياسها، فهناك متغيرات تتعلق بالبيئة أو السياق context أو المدخلات Inputs أو المخرجات Outputs وكذلك المتغيرات المتعلقة بالعمليات Processes.

ومن الجدير بالذكر أن البيانات قد يكون بعضها كمياً Quantitative كتلك التي يتم جمعها من خلال استمارات الملاحظة المنظمة والاستبيان أو المقابلة المقيدة.

بينما هناك معلومات أخرى تعتبر بيانات نوعية Qualitative لا تعتمد على مقاييس محددة الاستجابة. كالملاحظات الطبيعية، والمقابلات المعمقة يجري تحليل وتفسير نتائجها بناءً على خبرة المقوّم وتجاربه في العمل الذي يقوم بتقويمه وهي لا تعتمد في الغالب على المقاييس والأساليب الإحصائية. أما الأساليب الإحصائية الخاصة بتحليل البيانات الكمية فهي كثيرة ومتنوعة ومتفاوتة التعقيد منها (عرض البيانات في جداول خاصة لتخدم أغراض معينة، أو عرضها بجداول تكرارية وأشكال بيانية، وهناك التحليل المعتمد على الإحصاء الوصفي Descriptive Statistics مثل (مقاييس النزعة المركزية، التشتت، العلاقات وغيرها..) وهناك الإحصاءات الاستدلالية Inferential Statistics مثل اختبار "ت" t- test وتحليل التباين البسيط ANOVA وتحليل التباين المشترك ANCOVA وكلها تستخدم في اختبار الفرضيات الإحصائية وفقاً لنوع متغيرات البحث وبياناته.

المرحلة الثالثة: تقييم جودة التعليم:

بعد عملية تفريغ البيانات التي تم جمعها وتحليلها وفق نظام معين تجري مقارنة تلك المعلومات مع المعايير التي تم تحديدها مسبقاً وكما وردت في الصفحات السابقة ولكل عنصر من عناصر التقويم، وفي ضوء المقارنة بالمعايير يتم التأشير على نقاط الضعف والقوة والحكم على أداء المدرسة بشكل كلي وهو ما يمثل تحديد أو تقويم جودة الأداء المدرسي. ومن الجدير ذكره هنا أن المقوّم وبعد استعراض نتائج تحليل البيانات يحاول تفسير النتائج Interpretation الذي يعتمد على الفحص الدقيق لنتائج التحليل. وقد يكون التفسير متداخلاً مع تحليل البيانات بحيث يصبح مظهراً خاصاً من مظاهر عملية التحليل وليس عملية مستقلة. فتفسير البيانات هو البحث عن معنى لنتائج التقويم، حيث لا يقتصر فريق التقويم على وصف الواقع الراهن وإنما يحاول استخلاص توجهات ومؤشرات من البيانات التي قد تكون أحياناً متعارضة، وربط النتائج بالمتغيرات البيئية أو الداخلية لتفسير بعض النتائج الملفتة للانتباه والتي أسفر عنها التحليل الإحصائي.

كما أن فحص وتفسير نتائج البيانات النوعية المستمدة من المقابلات والملاحظات يزود فريق التقويم بمعلومات مفيدة فيما يتعلق بجوانب العمل المدرسي المختلفة المنظمة والاستبيانات الخاصة بآراء المعلمين والطلاب وأولياء أمورهم. وينبغي أن يستند تفسير نتائج تحليل البيانات على مؤشرات الأداء Performance Indications التي صممت أدوات القياس في ضوءها وبذلك يسهل تحديد وحصر إنجازات المدرسة والعوامل التي أسهمت فيها، وجودة أداء الطلاب ومدى كفاية المعلمين والإدارة المدرسية، وملائمة المباني والتجهيزات والمرافق وكفاءة استخدامها وتوظيفها، وفاعلية الاتصال بأولياء الأمور والمجتمع المحلي، وتحديد الفاقد في استخدام الإمكانيات المادية والموارد المالية، وتقدير الاحتياجات، وغير ذلك من المعلومات المفيدة التي يمكن أن تتضمنها البيانات.

المرحلة الرابعة: وضع أهداف واضحة، لمعالجة نقاط الضعف وتحديد أولويات التطوير:

بعد أن تتم عملية التقويم الذاتي وفق الخطوات المعروفة لأي عملية تقويم تلك التي تبدأ بتحديد أهداف التقويم ووضع خطته وتحديد أدوات جمع البيانات ثم عملية تحليل البيانات للتوصل إلى النتائج التي تتم مقارنتها بالمعايير المحددة لإصدار الأحكام وهي الخطوة النهائية في التقويم.. تأتي عملية وضع الخطط الخاصة بمعالجة نواحي الضعف وتعزيز جوانب القوة وتحديد أولويات التطوير.. وتتمثل قدرة المدرسة في هذه الخطوة، إذ من خلال إدراكها لطبيعة العمل واستناداً إلى نتائج التقويم تبدأ المدرسة بمناقشتها مع أعضاء هيئة التدريس لاقتراح أفضل السبل لمعالجة نواحي الضعف وتحديد البداية في أعمال التطوير.

فلو فرضنا أن من المؤشرات السلبية التي أفرزتها عملية التقويم هي ضعف الأساليب التدريسية التي يستخدمها المعلمون في التدريس، فما هو الأسلوب الذي توصي به إدارة المدرسة لمعالجة هذا الجانب.. هل العمل على إشراك المعلمين بدورات تدريبية مركزية، إما بإجراء دروس نموذجية داخل المدرسة بإشراف الموجه

التربوي والمعلم الأول؟ وإذا ما ظهر أن نسبة من الطلبة يواجهون صعوبة في حل تمارين تطبيقية في الرياضيات، فما هي الإجراءات التي يمكن اتخاذها لمعالجة هذه المشكلة؟ إعطاء محاضرات إضافية لهم في مادة الرياضيات، تكثيف المتابعة، تكليف الطلبة الأقوياء بمساعدتهم؟ أم تشجيع أساليب التعلم التعاوني للتخفيف من هذه الظاهرة؟ وهكذا يتم استعراض نقاط الضعف واحدة واحدة وترسم لها الخطط الكفيلة بتخفيفها أو الانتهاء منها ضمن فترة زمنية محددة، إن أساليب التطوير المقترح إجراؤها في المدرسة فيجب مناقشتها على صعيد جميع أعضاء الهيئة التدريسية والموجهين التربويين ومن لهم علاقة بعمل المدرسة وتحديد الأولويات للبدء بها ضمن سقف زمني محدد (أي تاريخ البدء وتوزيع العمل على فترات زمنية، ثم تحديد الوقت الملائم لإجراء تقويم لنتائج عملية التطوير.. وتحديد نوعه أي (تقويم تكويني أم تقويم نهائي).

المرحلة الخامسة: **تنفيذ خطط التطوير:**

بعد أن يتم وضع خطة التطوير التي تستهدف معالجة نقاط الضعف وتعزيز جوانب القوة، وبناءً على الخطة الزمنية التي حددت بداية التنفيذ وانتهائه تبدأ هذه المرحلة المهمة من مراحل التقويم وتتطلب القيام باجراءات عملية تعكس تنفيذ ما ورد في الخطة التي وضعت ويطالب الاشخاص الذين سيقومون بهذه المهمة بالعمل بتوفير المستلزمات المادية اللازمة للتطوير، وهنا لا بد من الذكر بأن بعض هذه المستلزمات كالدروس النموذجية أو تنظيم ورش عمل داخل المدرسة أو تشجيع التعلم التعاوني وغير ذلك قد يكون ضمن امكانات المدرسة، لكن اشراك بعض المعلمين في دورات مركزية، أو توفير الجوانب المالية اللازمة للتخفيف من مشكلة قائمة أو سد الشواغر الناشئة عن نقص المعلمين وغير ذلك قد لا تقع ضمن امكانات المدرسة لذلك فان ادارة المدرسة يمكن أن تستعين بالمديرية العامة للتربية في المنطقة لتوفير تلك المستلزمات.
وتعتبر هذه المرحلة من أهم الراحل لانها تمثل مرحلة الاجراء الفعلي لمعالجة نواحي القصور في العمل الذي ينعكس سلباً على الاداء العام للمدرسة.

المرحلة السادسة: الإشراف والمتابعة:

ويمكن أن يشترك في هذه المرحلة كل من إدارة المدرسة والمعلم الأول، ومشرفين من المديرية العامة في المنطقة التعليمية لوضع جدول زمني خاص بمتابعة سير عملية التنفيذ لخطط المدرسة في معالجة نقاط الضعف وفق الخطة المحددة سلفاً، ورفع تقارير دورية عنها، وكذلك الحال بالنسبة لأولويات التطوير للأداء المدرسي التي وضعت بعد الانتهاء من عملية التقويم، إن اقتراح مساهمة تمثيل المديرية العامة للتربية فيما يتعلق بالمتابعة يأتي على أساس أنها الجهة الرسمية العليا في المنطقة التي يهمها تطبيق الإجراءات المحددة في ضوء نتائج التقويم ومتابعتها من جهة، وتقديم العون والمساعدة للمدرسة، إذ قد لا تكون تلك الأمور ضمن إمكانات المدرسة لا سيما الأمور المادية أو المالية أو ما يتعلق بالجوانب الإدارية وغيرها.

كتابة التقرير الختامي:

يقوم فريق التقويم الذاتي بعد الانتهاء من جمع البيانات المطلوبة عن الجوانب التي خضعت لعملية التقويم وبعد تحليلها وتنظيمها وفقاً لأساليب كمية أو نوعية أو كلاهما يتم البدء بكتابة التقرير النهائي الذي يفترض أن يتضمن ثلاثة جوانب رئيسية هي:

1- وصف الوضع الراهن كما هو.
2- مقارنة الوضع الراهن بمعايير محددة مسبقاً لأجل إصدار أحكام موضوعية عن أداء المدرسة.
3- وأخيراً تحديد جوانب القوة من جهة وجوانب القصور التي تحتاج إلى التطوير والتحسين.

وقد حدد العديد من الباحثين مواصفات التقرير الختامي كالآتي:

1- أن يكون التقرير واقعياً وموضوعياً ومختصراً، أي أنه يعكس الواقع كما رآه الفريق وما جمعه من بيانات كمية ونوعية تشخص نقاط الضعف مثلما تشخص نقاط القوة.

2- يعكس أحكاماً واضحة عن كل عنصر من العناصر المدرسية التي تم تقويمها استناداً إلى معايير موضوعية وشواهد واقعية بعيدة عن التحيز الشخصي والأحكام المسبقة.
3- يربط خطة التطوير مع الأهداف العامة للمدرسة ونتائج التقويم.
4- أن يكتب بأسلوب واضح ومبسط يمكن فهمه بسهولة من قبل المسؤولين التربويين.
5- استخدام الأساليب الإحصائية المبسطة والابتعاد عن الأساليب المعقدة التي قد يصعب استيعابها من قبل المسؤول الإداري الذي يقدم له التقرير.
6- أن تكون التوصيات التي يتضمنها التقرير واقعية وذات طابع إجرائي يمكن الأخذ بها وتنفيذها من قبل المسؤول عن اتخاذ القرار.

أنواع تقارير التقويم:

تختلف تقارير التقويم الخاصة بالمدارس باختلاف الأطراف والجهات التي تهتم بنتائج التقويم واستخدامها، فالقيادات التربوية العليا بالوزارة على سبيل المثال أو المديريات العامة للتربية في المناطق التعليمية قد لا يكون لديها الوقت الكافي لقراءة تقرير تفصيلي فني طويل يتطرق إلى دقائق الأمور التي قد لا يكون عرضها مهماً في التقرير لا سيما أنها جاءت وفقاً لأدوات جمع البيانات المرفقة بالتقرير والتي يمكن الإطلاع عليها بسهولة من قبل المسؤول الإداري الأعلى الذي قد يحتاج إلى ما يسمى في التقويم (Executive Summary) اي الخلاصة التنفيذية التي تعكس نتائج التقويم في عبارات ومصطلحات مختصرة غير فنية نسبياً، ومجموعة من المقترحات والتوصيات الإجرائية التي تستند إلى ما توصلت إليه النتائج من تحديد جوانب القوة والضعف، والمشكلات التي تتطلب إجراءات مناسبة دون تقديم تفاصيل كثيرة ومملة أحياناً.

أما مديرو المدارس وهيئاتها التعليمية والفنية فإنها تحتاج إلى مزيد من التفاصيل تتضمن موجزاً وصفياً للمعلومات الأساسية المتعلقة بالمدرسة والإجراءات التي اتبعت في التقويم والقضايا التي برزت وتمت دراستها، مع

ذكر التوصيات التي انطلقت من الأحكام التي حددت لتقويم الحالة، وقد لا تنبثق هذه التوصيات أحياناً عن البيانات والنتائج، بل تنبثق من خبرات ومعرفة فريق التقويم المستندة إلى الواقع اليومي وهذا يكثر وجوده في تقارير (التقويم الذاتي) على وجه الخصوص، ويطرح العديد من العاملين في ميدان التقويم التربوي نماذج عديدة للتقرير الختامي لعل أوضحها من وجهة نظرنا ذلك الذي طرحه كل من وورثن وساندرز (Worthen & Sanders, 1973) نوجزه بالآتي:

القسم الأول: أهداف التقويم:

ويتضمن هذا القسم تحديد أهداف عملية التقويم لبرنامج أو مدرسة معينة، أي توفير إجابة عن سؤال فحواه لماذا نقوم بالتقويم وبأي أسلوب؟ ومن هم الأفراد الذين سيقدم لهم نتائج عملية التقويم؟

القسم الثاني: وصف البرنامج المقوّم (مدرسة، كلية، منهج ... إلخ).

وفي هذا القسم يتم الحديث عن مكونات البرنامج المقوم، وما الذي يجري فيه، وذلك من خلال استخلاص البيانات التي تم جمعها عن مكونات البرنامج سواء في عمليات التعلم، التعليم، القيادة المدرسية، ويتطرق إلى فلسفةوأهداف المؤسسة قيد التقويم، ويجري خلاله ايضاً التطرق إلى المعايير المستخدمة في المؤسسة لتحديد جوانب القوة والضعف في مجالات التعلم والتدريس، المعلم، الطالب والمادة التدريسية والمجالات الأخرى ويتضمن وصفاً للأدوات المستخدمة في عملية التقويم وخصائصها وفاعليتها وملائمتها لقياس ما وضعت لقياسه.

القسم الثالث: مخرجات البرنامج

ويتضمن التأكد من تأثير البرنامج المقوّم سواء على سلوك الطالب وملاحظة مدى التغير الحاصل فيه، أو حتى ملاحظة مدى التغير الذي تتركه المدرسة على المجتمع المحلي، وكذلك يتضمن كلفة البرنامج المالية، سواء ما يصرف على المستلزمات البشرية أو المادية.

القسم الرابع: العلاقات والمؤشرات

أي التوصل إلى تحديد مستوى العلاقة بين الأهداف المحددة للمؤسسة والملاحظات التي جمعت من خلالها البيانات عن الواقع الفعلي، أي إيجاد العلاقة بين الواقع المتحقق والأهداف المعلنة.

القسم الخامس: الأحكام القيمية

ويعني ذلك توضيح وتحديد قيمة البرنامج المقوّم المستند إلى بيانات يُفترض أن يبرهن المقوّم قيمتها وصدقها وما إذا كانت تحقق أهداف التقويم، ويُبرز هذا القسم أيضاً تحديد العلاقة بين أهداف البرنامج والحاجات الاجتماعية المتوقع إشباعها من قبل تلك المؤسسة أياً كانت (Worthen & Sanders, 1973)

أما علام (2003) فيطرح نموذجاً آخر لتقرير التقويم نطرح فيما يأتي ملخصاً لأهم عناصره:

1- **المقدمة**: وتشمل تقديماً موجزاً عن المبررات التي أُجري التقويم بناءً عليها وما هي أهدافه وماذا حققته المدرسة في المستوى التعليمي، ونوعية التعليم الذي تقدمه، وفاعلية توظيف موارد المدرسة، وبعض الإحصاءات المتعلقة بالطلبة، وكثافة الصفوف والمعلمين، ومعلومات أخرى وصفية عن المدرسة.

2- **وصف كيفية إجراء عملية التقويم**: ويشمل الوصف أعضاء الفريق الذي يقوم بالتقويم ومدة التقويم، والمواد الدراسية التي تمت ملاحظتها، وعينات من أعمال الطلاب وكيفية فحصها من قبل المعلمين. والمناقشات التي أجراها فريق التقويم مع الطلبة والمعلمين والإدارة المدرسية وتحليل نتائج استبيانات قدمت للآباء والطلاب قبل البدء بالتقويم. ووصفاً للأساليب والأدوات التي استخدمها الفريق في جمع البيانات وطبيعة مؤشرات الأداء التي استند إليها التقويم.

3- **نتائج تقويم المكونات المختلفة للنظام المدرسي**: ويعد هذا القسم من أهم الأقسام الرئيسية في التقرير حيث تشمل نتائج التقويم وتحديد مستويات

الأداء من خلال الأدلة الموضوعية وأحكام الفريق التي توصل لها في ضوء اجتماعاته ومناقشاته حول الموضوعات التالية: (مباني المدرسة، تنظيم المدرسة، المنهج، أساليب التدريس، الإدارة المدرسية، مستويات تحصيل الطلاب وإنجازاتهم واتجاهاتهم نحو المدرسة، النمو المتكامل للطلبة دينياً وأخلاقياً واجتماعياً وثقافياً من خلال المناهج والتدريس) وأخيراً علاقة المدرسة مع اولياء الأمور والمجتمع المحلي).

4- **تقديم التقرير للأطراف المعنية والإفادة من نتائج التقويم**: أي تحديد الجهات التي يفترض أن يوجه لها التقرير النهائي، و يمكن أن يقدم التقرير النهائي والملخص التنفيذي إلى المديرية العامة للتربية في المنطقة التعليمية، والقيادات التربوية في الوزارة، للإطلاع عليه، ويمكن أن يطلع أولياء الأمور على الملخص التنفيذي من خلال إجتماعات مجالس الاباء والمعلمين.

متابعة التقرير:

تتم متابعة نتائج التقويم من قبل إدارة المدرسة للتحقق من تنفيذ المقترحات والتوصيات التي تضمنها التقرير، ومتابعة خطة التقرير وفق الجدول الزمني الذي يُفترض أن يحدده فريق التقويم الذاتي وتحديد مستوى التنفيذ أو معالجة العقبات التي قد تعترض سبيله، وأن تتم متابعة تنفيذ خطة التطوير مع جميع أفراد الملاك الإداري والفني والتعليمي في المدرسة لأجل أن يعرف كل عضو ما عليه أن يعمله لتنفيذ مفردات الخطة ولضمان تحقيق التطوير والتحسين المطلوبين.

مراجع الفصل الثاني

1- الزاملي، علي عبد جاسم (2006). محاضرات في التقويم الذاتي. القيت على العاملين في وحدات التقويم التربوي بوزارة التربية والتعليم. سلطنة عمان: مسقط.

2- الزوبعي، عبد الجليل؛ والعاني، نزار (1980). رأي في تطوير القياس والتقويم. مجلة العلوم التربوية والنفسية. 5، 65 - 78.

3- عودة، أحمد (2002). القياس والتقويم في العملية التدريسية، إربد: دار الأمل.

4- علام، صلاح الدين محمود (2003). التقويم التربوي المؤسسي: أسسه ومنهجياته وتطبيقاته في تقويم المدارس. القاهرة: دار الفكر العربي.

5- سعادة، جودت أحمد، وإبراهيم، عبد الله محمد (1997). المنهج المدرسي في القرن الحادي والعشرين. الكويت: مكتبة الفلاح للنشر والتوزيع.

6- وزارة التربية والتعليم، (2004). دليل مشروع تقويم الأداء المدرسي وتطويره. مسقط: دائرة تطوير الأداء المدرسي.

7- A- Cook T, D, & Aeichardt C, S. (E.d) (1979). **Qualitative and Quantitative Methods in Evaluation Reasearch**. California: SAGE Publications, Inc.

7B- Cooley, w.w. of Lohnes, P.R (1976). **Evaluation Research in Education**. New York: Irvington Publishers, Inc.

8- Fitzpatrik, J., Sanders, J., & Worthen, B. (2004). **Program Evaluation: Alternative Approaches and Practical Giuidines**. New York: Allyn & Bacon.

9- Guba, E. G., & Lincoln Y. S. (1981). **Effective Evaluation**. California: Jossey - Bass Publishers.

10- Guba, E. G., and Stufflebeam, D. L. (1968). **Evaluation: The Process of Stimulating, aiding and abetting insighiful action**. 2nd phi Delta kappa National symposium for professors of Educational Research colorado.

11- Millman, J. (E.d) (1981). **Handbook of Teacher Evaluation**. California: SAGE Publications, Inc.

12- **Self Evaluation and the New Common Inspection Framework. Aguide for schools** (2004). http: // www.ofsted.gov.uk.8/1/2006.

13- **Shool Self Evaluation, Guidance for Rochdale Schools** (2004). http:// www.rochdale. gov.uk. 8/1/2006.

14- Stake, R. E. (1967). The Countenance of Educational Evaluation. **Teacher College Records**, 68, 523- 540.

15- Stufflebeam, D. L, & Shinkfield, A. J. (1985). **Systematic Evaluation**. Boston: Kluwer nijhoff Publishing.

16- Wiles, K. (1972). **Changing Perspectives inEducational Research**. Ohio: A. Jones.

17- Worthen, B. R., & Sanders J. R. (1973). **Educational Evaluation: Theory and Practice.** California: Wadsworth Publishimg Company, Inc.

18- Worthen, B.R., & Sanders, J.R. (1987). **Educational Evaluation: Alternative Approaches and Practical Guidelines**. N.Y: Longman.

الفصل الثالث

تقويم البرامج التربوية باستخدام النماذج

أولاً: النماذج المعتمدة على الأهداف

ثانياً: النماذج التحكيمية

ثالثاً: النماذج المساعدة على اتخاذ القرار

رابعاً: النماذج النوعية في التقويم

3

الفصل الثالث

تقويم البرامج التربوية بإستخدام النماذج

قبل استعراض أنواع النماذج الشائعة لتقويم البرامج التربوية، لابدً وأن نوضح للقارئ ما المقصود بالبرنامج التربوي أو التعليمي الذي يُعرف بأنه: مجموعة من النشاطات المستخدمة للتعامل مع مشكلة معينة أو تجديد ممارسة قائمة، أو إدخال تطويرٍ معين على العملية التعليمية المستخدمة. (الدوسري، 2000) والبرنامج التربوي قد يتكون من عددٍ من النشاطات أو المشاريع المنفصلة ذات الأغراض المحددة التي تستخدم لتحقيق أهداف ذلك البرنامج، فإدخال منهج جديد لتعليم الرياضيات الحديثة مثلاً أو تدريب المعلمين على صياغة واستخدام الاختبارات الموضوعية يمكن أن تمثل أنواعاً من البرامج التعليمية.

والبرنامج التربوي بشكل عام أو التعليمي بشكل خاص قد يكون (وحدة قائمة بذاتها) وله أهدافه ومصادره، ويحدد الأشخاص الذين سيستفيدون منه، ويتكون البرنامج التربوي في الغالب من عناصر ومكونات مثل المدخلات والعمليات والمخرجات، وقد يتضمن في كثير من الأحيان (تغذية راجعه Feedback) تنفع في تجديد أو تطوير ذلك البرنامج. وللبرنامج بيئة Environ-ment تتمثل في خلفية البرنامج والتطور التأريخي له، وتحديد المسؤولين والمساهمين في ظهور فكرة ذلك البرنامج والأسباب التي أدت إلى إقامته والمشكلات التي وُجدَ البرنامج لمعالجتها، كما تتناول سمات المجتمع الذي يخدمه البرنامج سواء كانت سمات ديموغرافية أو مهنية أو غيرها.

أما مدخلات البرنامج Inputs فتتمثل في أهدافه وكيفية صياغتها والمصادر التي اشتقت منها، وتحديد المسؤول عن صياغتها وعلاقتها بالحاجات الفعلية للفئات التي يستهدفها البرنامج، وهل تمت الاستعانة ببعضهم للمشاركة في صياغة تلك الأهداف. وتشمل المدخلات أيضاً توفير معلومات هامة عن

المعلمين والإداريين والفنيين العاملين في البرنامج والطلبة ومستويات تأهيلهم والخبرات التعليمية التي يمتلكونها، كما تشمل معلومات عن المصادر المادية من أجهزة ومعدات ومواد دراسية وغيرها، يضاف إلى ذلك أن البيئة التنظيمية والتركيب الإداري تُمثل أجزاء مهمة من المدخلات.

أما العمليات Processes فيقصد بها الأنشطة والفعاليات المتمثلة بعمل البرنامج ومنها أساليب التدريس والتدريب والأنشطة الصفية واللاصفية، وأساليب التعامل والتفاعل مع الأفراد المستفيدين سواء كانوا طلبةً أو عاملين، وقد يشمل أيضاً أولياء أمور الطلبة.

أما المخرجات Outputs فيقصد بها النتاجات النهائية للبرنامج المتمثلة بمدى تحقق الأهداف الأساسية التي أقيم البرنامج لتحقيقها، وهي بالتأكيد تمثل الأدلة والشواهد على نجاح أو فشل البرنامج.

ومن المعروف والمفيد أن يخضع أي برنامج كله أو أجزاء منه لعملية التقويم وبشكل دوري إذا ما أُريد له النجاح والتطوير، وهناك استراتيجيات ونماذج متنوعة تستخدم في عملية التقويم التربوي. ويقصد باستراتيجية التقويم Evaluation Strategy الطريقة التي يتم اتباعها في عملية التقويم ابتداءً من مرحلة التخطيط لها وحتى تنفيذها والوصول إلى نتائجها، أي أنها تشكل الأسلوب أو الخطة التي يتم اتباعها للتوصل إلى قرارات تستند إلى معلومات صادقة وموضوعية عن الشيء المراد تقويمه. (علام، 2003).

أما نماذج التقويم Evaluation Model فهي تصورات فكرية نظرية للخطوات والإجراءات اللازمة للتوصل إلى قرار حول البرنامج التربوي المراد تقويمه. (الدوسري 2000) وقد يمتد مفهوم النموذج ليشمل الفلسفه والمنطلقات النظرية التي يُبنى تسلسل العملية في ضوئها مع توضيح مفصل للخطوات والمراحل والعلاقات القائمة بينها. (الشبلي، 1984).

وقد ظهرت بعض النماذج التقويمية في الميدان التربوي منذُ الأربعينات من

القرن الماضي لكنها ازدادت عدداً وتنوعاً بين العقدين السادس والثامن من القرن المذكور، حيث يُشير ورثن وساندرز (Worthen and Sanders, 1987) أنه بين عامي 1967 و 1987 تمَّ تطوير ما يَزيد عن خمسين نموذجاً مختلفاً للتقويم عندما ظهرت الحاجة الماسة إلى إجراء العديد من الدراسات التقويمية المنظمة وأدرَكَ الباحثون أهمية الاستناد إلى تصورات أو أُطر فكرية تُفيد في توجيه الممارسات الميدانية. وقد تعرضت هذه النماذج إلى إعادة صياغة وتعديلات نتيجة عاملين هما: الانتقادات التي وجهت لها من قبل باحثين آخرين، وما نشأ من ملاحظات أو صعوبات بعد تطبيقها في الميدان، ولذلك نرى أن بعض هذه النماذج بقي يعبِّر عن وجهات نظر فلسفية واسعة ويعكس توجهات اجتماعية معينة لدى مصمميها. ويرى ستيك (Stake 1991) إن ما توضحه أدبيات التقويم ليست نماذج بل (مداخل وقناعات) تعكس اهتمامات المنظرين وتعبر عن توجهاتهم وتؤكد وجهات نَظرٍ مُعينة. أما ورثن وساندرز (Worthen and Sanders 1973) فيرى أن قيمة التنبؤ في بعض النماذج التقويمية ضعيفة ولم يجر التحقق من صدقها بشكل تطبيقي عملي، ولذلك فهي لا تفي بمتطلبات النموذج العلمي رغم فائدتها في توجيه وإرشاد القائمين بالتقويم عند ممارستهم لهذه العملية. ويُشير علام (2003) إلى أنه رغم أن العاملين في ميدان التقويم يُشيرون أحياناً إلى تصميمات تقويمية تَستند إلى نموذج معين، غير أنهم لا يوضحون بالضبط الإجراءات الفعلية المستمدة من النموذج الذي استخدموه في الميدان العملي، لذلك يمكن القول أن الكثير من نماذج التقويم ليست توصيفية Prescriptive وإنما وصفية Descriptive وتفتقر إلى الإجرائية Operational مما يجعلها عرضةً لتفسيرات مُتباينة عند التطبيق في السياقات الميدانية، وفي العديد من تلك النماذج لابد للمقوِّم أن يتعرف بعمق على الإطار الفلسفي الفكري للنموذج الذي يُريد اختياره ومدى ملاءمته للبرنامج أو النشاط المطلوب تقويمه، وأن تكون لديه القدرة على تحديد نقاط ضعف وجوانب قوة النموذج المختار، ولذلك يمكن القول أن تنوع النماذج لا يعني

بالضرورة اعتماد واحدٌ منها باعتباره النموذج المفضل الذي يمكن استخدامه في جميع الأحيان بل أن النماذج تمثل أَساليب مقترحة من قبل باحثين معينين تهدف إلى توفير معلومات مهمة عن الكيفية التي سَيقوَّم بها برنامج ما متضمناً الطرق والإجراءات التي يُفترض أن يقوم بها الباحث أثناء عملية التقويم.

ونظراً لاختلاف الظروف المحيطة بكل برنامج تربوي سواء في مدخلاته أو عملياته أو مخرجاته فليس هناك نموذجاً واحد يصلح لتقويم كل البرامج، بل أن بعض النماذج والاستراتيجيات التي تناسب برنامج معين قد لا تناسب برنامجاً آخر، إذ يعتمد ذلك من الجهة الأخرى على طبيعة النموذج والغرض من التقويم ونوع القرار المطلوب اتخاذه بعد عملية التقويم. ونُقدم في الصفحات القادمة عرضاً لأهم النماذج التقويمية الشائعة في الميدان التربوي من أجل تزويد القارئ بالأسس النظرية والتطبيقية لتلك النماذج لغرض الاستفادة منها في عملية تقويم البرامج والنشاطات التربوية.

تصنيف النماذج: Classification of Evaluation Models

عندما نتصفح الأدبيات الخاصة بموضوع التقويم نُلاحظ أَنماطاً متنوعة من التصنيفات للنماذج أو الأساليب المقترحة لإجراء التقويم، ومن أجل تبسيط ذلك إرتأينا تقديم التصنيف التالي، والذي نحاول فيه ربط النماذج ببعضها بناءً على تشابه في الغرض والوسائل المستخدمة فيها، ثم نستعرض خصائص نماذج مختارة منها بناءً على سهولة وشيوع استخدامها وشهرتها الواسعة في مجال التقويم:

أولاً - النماذج المعتمدة على الأهداف. (النماذج الهدفية) ومنها:

أ. نموذج تايلر Tyler

ب. نموذج متفيسل ومايكل Metfessel and Michael

جـ. نموذج هاموند Hammond

د. نموذج بروفس Provus

ثانياً: النماذج التحكيمية (التي تعتمد على محك خارجي) ومنها:

أ. نموذج سكريفن (التقويم المتحرر من الأهداف (Goal Free Evaluation)

ب. نموذج ستيك (الهيئة الكلية Countenance)

جـ. نموذج التقويم المتجاوب لستيك (Responsive Evaluation)

ثالثاً: النماذج المساعدة على اتخاذ القرارات (نماذج القرارات) ومنها:

أ. نموذج مركز دراسات التقويم بجامعة كاليفورنيا في لوس أنجلس.
Center for the Study of Evaluation.

ب. نموذج ستفلبيم. (CIPP) Stufflebeam Model

جـ. نموذج التقويم المرتكز على الخبراء Experts - Oriented Model

رابعاً: النماذج النوعية في التقويم Qualitative Models in Evaluation **ومنها:**

أ. التقويم التنويري Illuminative Evaluation

ب. نموذج التقويم الطبيعي Naturalistic Evaliation

أولاً: النماذج المعتمدة على الأهداف. **Objective - Oriented Model.**

ومن أهم هذه النماذج الآتي:

أ. نموذج تايلر **Tyler Model**

يَعود الفضل في ظهور هذا النموذج إلى الخبير التربوي المشهور (رالف تايلر Ralf Tyler) الذي رَكَّزَ دراساته على أهمية الأهداف وتصنيفها في العملية التربوية، وترأس مجموعة دراسة الثمان سنوات التي صدر تقريرها عام 1942م والتي صممت لاختبار مدى فعالية المناهج واستراتيجيات التدريس المستخدمة في المدارس الأمريكية في تلك الفترة. لقد كان تركيز هذا النموذج منصباً على الأهداف كنقطة انطلاق في عملية التقويم، وأن الهدف الأساسي لهذه العملية هو تحديد مدى تحقيق المنهج أو البرنامج قيد الدراسة للأهداف المرسومة له. ولذلك يركز هذا النموذج على جانبين مهمين ويمكن تحديدهما كمياً وهما: مخرجات العمل ونتائجه في الجماعة المستهدفة والثاني تحديد

التباين بين الأهداف المعلنة والنتائج الفعلية. والبرنامج المقوَّم قد يكون برنامج تدريبي، برنامج علاجي، تعليمي. أما الجماعة المستهدفة فهي الجماعة التي يُتوَقع من البرنامج أو أهدافه إحداث تغيرات في سلوكها من خلال تحديد مستوى تعلمها مثل الطلبة، الأميين، الموظفين..الخ).

أما الخطوات الرئيسية التي يسير عليها النموذج في عملية التقويم، فتتلخص كالآتي:

1- تحديد الأهداف.
2- تصنيف الأهداف.
3- تعريف الأهداف إجرائياً.
4- تحديد مواقف يمكن أن تكون مؤشراً لتحقق الأهداف.
5- تطوير أدوات قياس أو اختيار أدوات جاهزة بعد التأكد من موضوعيتها وقياسها لمواقف التقويم كأن تكون (مقاييس تقدير، استمارة ملاحظة، مقابلة، أو مقاييس إتجاهات).
6- جمع بيانات ومعلومات عن تلك المواقف أو ما يُعرف بمؤشرات الأداء.
7- مقارنة الأداء فعلياً بالأهداف المعلنة والمحددة مسبقاً.

وتعتبر الأهداف غير المتحققة نقصاً في البرنامج المقوَّم، أما الأهداف المتحققة فتعتبر نجاحاً للبرنامج. وأهم ما يُميز هذا النموذج بساطته وسهولة فهمه وتطبيقه، إذ عادةً ما يؤدي إلى معلومات واضحة ومقبولة من قبل المربين، كما أنه يدفع المسؤولين والقائمين على البرنامج إلى التركيز على الدقة والموضوعية في صياغة أهداف محددة لبرامجهم يمكن متابعة تنفيذها من جهة وسهولة قياسها من جهة أخرى. كما أن ذلك يمكن أن يؤدي من جهة أخرى إلى تنوع المقاييس التربوية المستخدمة في المؤسسات التربوية عبر اختبارات الورقة والقلم (Sanders & Sachse, 1977).

ورغم الميزات الحسنة لهذا النموذج، فقد وجهت بعض الانتقادات له ومنها أن نتائج التقويم تأتي متأخرة حتى نهاية البرنامج مما يحد من الاستفادة

منها في تحسين البرنامج وتطويره. يضاف إلى ذلك الصعوبات التي تواجه المقوَّم في كيفية تحديد الأهداف وإزالة التشابك بين بعضها البعض من جهة وتخفيض نسبة العمومية فيها أو الغموض التي تؤدي من جهة أخرى إلى صعوبة إجراء المقارنة بين الأداء وبينها بشكل دقيق وواضح لتحديد الفرق بينهما أي الحكم على مدى التباين بين الأهداف ومؤشرات الأداء الحقيقي، كما أن هذا النموذج قد يهمل بعض المخرجات الهامة التي لا تغطيها الأهداف المحددة (الصارمي، 2003).

ب. نموذج هاموند Hammond Model

ويُعتبر (هاموند) من الباحثين الذين ساروا على هدى (تايلر) في اعتماد عملية التقويم على أهداف البرنامج ويرى أن تقويم أي برنامج يعتمد على تحديد فاعلية ذلك البرنامج في تحقيق الأهداف المرسومة له. ويتمثل نموذج هاموند في اتباع المقوِّم لاستراتيجية تتكون من عدة خطوات يُرمز لها بالحروف (EPIC) وهي كالآتي:

1- تحديد الجوانب المراد تقويمها في برنامجٍ معين.
2- تحديد الأهداف العامة للبرنامج المقوَّم وتصنيفها.
3- تحليل الأهداف العامة إلى أهداف سلوكية يمكن قياسها.
4- تحديد المواقف التي يمكن أن تُلاحظ فيها الأهداف التي وضعت بشكل إجرائي (سلوكي).
5- تصميم أو اختيار أدوات موضوعية لقياس تلك الأهداف.
6- جمع البيانات الخاصة بالأهداف من خلال تطبيق الأدوات المشار إليها في الخطوة السابقة.
7- إجراء مقارنة بين البيانات التي يتم جمعها والأهداف التي تم تحديدها.
8-من خلال إجراء المقارنة المشار إليها في الخطوة السابقة سيتوصل المقوِّم إلى تحديد ما تم تحقيقه من أهداف البرنامج، فإذا ما تبين عدم تحقق الأهداف، أو تحقق القليل منها، يمكن الحكم على أن هناك ضعفاً أو قصوراً في البرنامج، أما إذا لاحظ المقوِّم أن معظم أهداف البرنامج قد تم تحقيقها

فسيتم الحكم عليه بأنه برنامج ناجح، ومن الجدير بالذكر أن هذا النموذج يهتم بالتغذية الراجعة Feed Back التي تفيد في إعادة صياغة أو تطوير أو استبدال وسائل خاصة بالبرنامج قد ينتج عنها تحسين ذلك البرنامج.

ومن الجدير بالذكر أن (هاموند) لم يبتعد عن الأسلوب الذي اعتمده (تايلر) في اعتبار الأهداف كمحكات للحكم على البرنامج، لكنه أضاف إلى ذلك التفاعل الذي يحدث في البيئة التربوية بين عناصر ومتغيرات منها متغيرات سلوكية، ومتغيرات تعليمية ومتغيرات مؤسسية وأخذها بنظر الاعتبار للتعرف على مدى تأثيرها في تحقيق النواتج المرجوه، كما اهتم بتحديد وتحليل التفاعل بين المتغيرات التعليمية والمؤسسية لتحديد تأثيرها على سلوك المتعلمين. ويشير كل من ورثن وساندرز (1987) إلى أن هاموند اقترح ضرورة تدريب الأفراد في البيئة المحلية التي يوجد فيها البرنامج بواسطة مدربين خارجيين متخصصين في تقويم البرامج ليكونوا قادرين على إجراء ذلك التقويم، حيث تعتبر هذه الخطوة استراتيجية متميزة بإشراكها أفراد محليين في عملية التقويم.

وفيما يلي مثال تطبيقي لخطوات هذا النموذج:

لو فرضنا إننا نريد تقويم برنامج (إثراء القدرة القرائية لأطفال المرحلة الابتدائية) بموجب خطوات نموذج هاموند فإنه يتوجب علينا القيام بالآتي:

1- تحديد جوانب المراد تقويمها في البرنامج وتتمثل في التأكد من أن الأطفال تحسنت قدرتهم القرائية وأصبح بإمكانهم قراءة كلمات مكتوبة بشكلٍ منفرد أو ضمن عبارات محددة الطول على سبورة الصف أو في سياق مطبوع.
2- يحاول المقوِّم التأكد من أن الأهداف العامة للبرنامج محددة بشكلٍ واضح بحيث تتطابق مع الغرض المذكور في الفقرة الأولى.
3- تحويل الأهداف العامة إلى أهداف سلوكية يُمكن قياسها.

بما أن الأهداف في كثير من الأحيان تصاغ بعبارات واسعة يصعب قياسها فإن على المقوِّم أن يحول تلكَ العبارات إلى صورٍ إجرائية سلوكية يمكن قياسها بسهولة كأن يضعها بالشكل الآتي:

- يستطيع التلميذ قراءة خمس كلمات في ثانية واحدة.

- يستطيع الطالب أن يربط بين كل كلمة مكتوبة مع الصورة الدالة عليها من بين مجموعة من الصور.
- يميز من بين عدة كلمات الكلمة التي تدل على صورةِ حيوان معين.
- يعطي معنى كلمة معينة بأسلوبه الخاص.
- يدخل كلمة معينة في جملة مفيدة ومفهومة.

4- تحديد المواقف التي يمكن أن تُلاحظ فيها الأهداف السلوكية: ويمكن تنفيذ ذلك من خلال القراءة في كتاب، أو على السبورة، نطق الحروف بشكل صحيح، التمييز بين الصور والكلمات الدالة عليها في كتاب أو في وسيلة تعلمية معلقه على سبورة الصف.

5- تصميم أو انتقاء أداة للقياس: هل يستخدم المقوِّم اختباراً قرائياً جاهزاً ومعروفاً؟ أم يطور مقياساً خاصاً به لهذا الغرض. وهل الاختبار شفوي أم كتابي.. الخ.

6- جمع البيانات الخاصة بالأهداف: يتم جمع البيانات بشكل تفصيلي حسب الخطوة السابقة وتحديد عدد الإجابات الصحيحة وعدد الإجابات الخاطئة ونسبها لمجموعات التلاميذ، وفقاً لقواعد الاختبارات التي تمَّ تطبيقها.

7- مقارنة نتائج قياس الأداء بالمعايير التي تم تحديدها في بداية عملية التقويم باستخدام أساليب إحصائية مبسطة (نسب مئوية، متوسطات حسابية، انحرافات معيارية، ..الخ) للتأكد من تحقق أو عدم تحقق الأهداف المحددة مسبقاً للبرنامج، مثال ذلك:

إذا كانت درجة اتقان الطلبة للمهارات القرائية كالتالي:

أقل من 50%	تحدد بأنها نسبة ضعيفة
60-51	مقبولة
70-61	متوسطة
80-71	جيدة
90-81	جيدة جداً
100-91	ممتازة

8- يتم الحكم على كل هدف من أهداف البرنامج بمدى تحققه بمستوىً معين، ويتم الحكم على البرنامجً بشكل عام استناداً إلى تحقق جميع أهدافه المحددة.

جـ. نموذج متفيسل ومايكل **Metfessel & Michael Model**

تتألف الاستراتيجية التي اقترحها كل من متفيسل ومايكل من ثمان خطوات إجرائية تشبه إلى حد ما تصميم هاموند لكنها تتميز عنه في اهتمامها بمقاييس محكيه متعددة Multiple Criterion Measures وهذه الخطوات هي:

1) ضرورة إشراك أفراد المجتمع المهتمين بالبرنامج المقوم في عملية التقويم سواء بصورة مباشرة أو غير مباشرة.
2) تحديد الأوليات التي يسعى البرنامج لتحقيقها وفقاً لحاجات المجتمع.
3) بناء نموذج للأهداف الرئيسية للبرنامج يشتمل على تنظيم هرمي حسب عمومية تلك الأهداف أو نوعيتها، ابتداءً بالغرض العام ثم الأهداف العامة التي تخدم تحقيق الغرض العام، وتحليل تلك الأهداف إلى أهداف إجرائية (سلوكية) يمكن قياسها بسهولة بأدوات مصممة لهذا الغرض، تم تحديد المحكات التي في ضوئها يمكن الحكم على النواتج الخاصة بالبرنامج ومدى مطابقتها لاحتياجات المجتمع، ومصادر التمويل للبرنامج وأساليب إعداد الأفراد في البرنامج.
4) تصميم أو اختيار أدوات قياس متنوعة تتميز بالخصائص السيكومترية من صدق وثبات لكي تستخدم في جمع البيانات وتحليلها لتحديد مدى تحقيق البرنامج للأهداف التي تمَّ تحديدها بصورة إجرائية مسبقاً.
5) تطبيق الأدوات بشكل دوري.
6) تحليل البيانات وفقاً لأساليب إحصائية مناسبة.
7) تفسير البيانات في ضوء معايير أو محكات معينة مثل (مدى التقدم الحاصل في جوانب معينه، أو فاعلية البرنامج بشكل عام).
8) وضع التوصيات والمقترحات المتعلقة بالبرنامج في ضوء تفسير البيانات.

ويقترح واضعا النموذج ضرورة أن تقدم تغذيه راجعه بعد كل خطوة ليتم بموجبها إجراء التعديلات اللازمة في أساليب التقويم بما يشمل حتى أدوات التقويم المستخدمة، لذلك قدّما قائمة بالأدوات المحكية التي يمكن أن يسترشد

بها المقوِّم (علام، 2003). ولو كان البرنامج التربوي المقوَّم نشاطات تعليميه فإن النموذج يقتضي التأكيد على جمع معلومات عن مؤشرات هامة في البرنامج منها:

1) مؤشرات التغيير الذهني والوجداني في سلوك الطالب (معرفي، وجداني) وفق معايير واختبارات مقننه، أو وفق معايير رسميه وغير رسميه يطورها المعلمون مثل: الملاحظات المباشرة والاختبارات المدرسية، وكشوف حضور وغياب الطلبة، والمشاركة في نشاطات الفصل وإنجاز الواجبات.
2) مؤشرات التغيير الذهني والوجداني في سلوك المعلمين والعاملين في المدرسة.
3) مؤشرات التغيير الذهني والوجداني في سلوك المجتمع والمحيطين بالبرنامج.

د. نموذج بروفس (تقويم الاختلاف أو التعارض) **Provus Discrepancy Evaluation Model**

تعتمد فكرة هذا النموذج على الموازنة (المقارنة) بين الأداء الفعلي والأداء المطلوب، لذلك يركز على الاختلاف بين هذين الجانبين وفقاً لمستويات موضوعية يتم تحديدها قبل البدء في تنفيذ البرنامج، ولذلك يطلق عليه "نموذج التقويم المبني على الاختلاف بين الأداء الفعلي والمستويات المطلوبة"، وكنتيجةً لعملية التقويم يدعو واضع النموذج إلى استخدام نتائج التقويم بعد تحديد الاختلاف، إلى تغيير الأداء الفعلي أو إلى تغيير مستويات الأداء المطلوبة وحسب الظروف والمتغيرات المحيطة بالبرنامج. يُضاف إلى ذلك أن الاختلاف الذي يقصده (بروفس) قد يكون موجباً أو سالباً، فإذا زاد الأداء عن المستوى المطلوب أو المحدد مسبقاً من قبل المقوِّم دَلَّ ذلك على أن الاختلاف موجباً، أما إذا قلَّ عن المستوى المحدد دَلَّ ذلك على أن الاختلاف سالب، وعند ذاك يجب التعامل مع البرنامج وفقاً لأحد الأساليب التالية:

(1) يمكن إعادة تشكيل أو صياغة المستويات التي قد لا تكون واقعية أو ملائمة لطبيعة البرنامج.

(2) إعادة النظر في الضبط الإداري للبرنامج وزيادته للتحكم بصورة أفضل في الأداء.
(3) يمكن التوصية باتخاذ قرار بإلغاء البرنامج إذا تبين أن الاختلاف كبير بين الأداء الفعلي والأداء المطلوب.

ومما تجدر الإشارة إليه هُنا أن هذا النموذج يعتمد فيه تقويم البرنامج والحكم عليه استناداً إلى المستوى Standard سواء كان المستوى كمياً أو نوعياً، وبدونه تصبح عملية التقويم غير ممكنة. ويُشير واضع التصميم إلى أن تحديد المستوى يتضمن ثلاث عمليات تؤثر إيجابياً في تخطيط البرنامج وهي:
- العملية الأولى: وتتضمن جمع المعلومات اللازمة لوصف وتحديد مدخلات وعمليات ومخرجات البرنامج.
- العملية الثانية: وتتضمن التوصل إلى اتفاق عمل عن طريق التوفيق بين الآراء أو القيم المتعارضة دون قَسر أو إلزام.
- العملية الثالثة: وتتضمن تقدير كِفاية تصميم البرنامج من الناحية الفنية بحيث يمكن الاعتماد عليها كأساس في تقويم البرنامج.

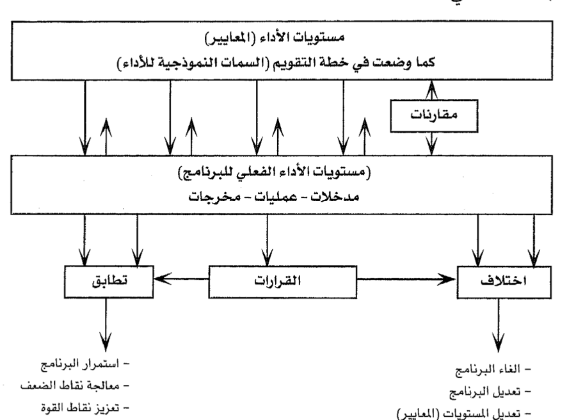
إجراءات التقويم وفقاً لهذا النموذج:

يقترح بروفس عدة خطوات لتقويم البرنامج وفقاً لهذا النموذج هي: تصميم خطة التقويم، جمع المعلومات عن الأداء طبقاً للخطة، وتقرير الاختلاف الذي يعتمد على موازنة الأداء بالمستوى وأخيراً مواجهة هذا الاختلاف بإحدى الحلول المقترحة سابقاً.

ويورد (الدوسري، 2000) إن مشكلة تحديد المعايير أو المستويات الخاصة بالتقويم تعتبر المشكلة الأولى بالنسبة للمقوِّم، ولتعقد هذا الموضوع يلجأ بعض المقومين إلى عدة أساليب لتجزئتهَ إلى عناصره الأساسية لتحديد السمات والخصائص التي يمكن أن تمثل المعايير المرغوبة. فلو طلبَ إلى مقوِّم ما تقويم برنامج تعليمي مثل (تطوير نظام الإشراف التربوي في التعليم العام) حيث يجري التركيز على المكونات الرئيسية للبرنامج مثل (المدخلات، العمليات والمخرجات)، وتطوير معايير خاصة لكل مكون من تلك المكونات الأساسية والفرعية. ففي مجال المدخلات يمكن النظر في أهداف البرنامج ومدى

وضوحها وسهولة تحقيقها مثل (العدد المطلوب من المشرفين)، وفي العمليات يمكن النظر إلى الإجراءات التي تتم أثناء تنفيذ عمل الإشراف التربوي كالزيارات الميدانية للمدارس والأسلوب المتبع في تقويم المعلم، وإجراء مقابلات بين المقوِّم وعينة من المعلمين للتأكد من مدى تحقق الأهداف المرسومة لعملية الإشراف، وتحديد اتجاهاتهم نحو عملية الإشراف التربوي.

وفي مجال المخرجات يمكن النظر إلى مدى تأثير عملية الإشراف التربوي في تطوير أساليب التدريس والتقويم وما ينعكس منها على التحصيل الدراسي للطلاب. ومن خلال المعلومات التي يجمعها المقوّم من القنوات المذكورة يمكن تطوير معايير يحتكم إليها لتحديد مدى الاختلاف والتقارب بين الأداء والمعايير، فإن لاحظ المقوّم فروقاً كبيرة بينهما اقترح على صاحب القرار اتخاذ إجراءات تطويرية لردم الهوةِ وإيجاد التطابق بينهما. ويمكن تمثيل هذا النموذج بالمخطط التالي:



شكل (1-3) مخطط لنموذج بروفس

ثانياً: النماذج التحكيمية (التي تعتمد على محكٍ خارجي)

تعتمد النماذج التي سبق ذكرها في الصفحات السابقة (تايلر وآخرون) على محكات داخلية هي (الأهداف) الخاصة بالبرنامج، وقد انتُقدَت لأنها تعتمد علىً (محكات العمليات Process Criteria) فقط دونَ محكات خارجية، لذلك ظهرت هذه النماذج الجديدة المعتمدة على محكات خارجية External Criteria، ومن أكثر تلك النماذج شيوعاً الآتي:

أ. **نموذج سكريفن Scriven**

وجَهَ سكريفن إنتقادات شديدة إلى منهجية التقويم ونماذجه المختلفة، وكان لهذه الانتقادات أثر واضح في تعديل بعض مفاهيم التقويم وتطوير المنهجية التي يتبعها. ولعل أهم هذه الانتقادات قضية التركيز على الأهداف. حيث أوضح أن المهم هو ليس تقويم أهداف البرنامج بل التركيز على نوعية تلك الأهداف والحكم على جدارتها، إذ لا فائدة من تقويم أهداف لا تستحق أن نسعى إلى تحقيقها، لذلك يتوجب على المقوِّم أن يحدد مدى جدارتها.

لقد قدّم (سكريفن) الكثير من المفاهيم التي توجه لخدمة المستهلك أو المنتفع من البرنامج التعليمي حيث قدَّم نوعين من التقويم هما:

التقويم البنائي Formative Evaluation والتقويم الختامي Summative Evaluation، كما أكد على ضرورة تقويم الأثر أو مدى تأثير البرنامج من خلال التعرف على نواتج البرنامج سواء كانت مقصودة أم غير مقصودة، مباشرة أو غير مباشرة، إضافةً إلى تقويم كل ما يجري داخل البرنامج من تفاعلات ونشاطات.

كما أكد سكريفن على أهمية (التقويم المقارن) أي أن يجري تقويم البرنامج مقارنةً بغيرهِ من البرامج للمساعدة في اختيار البدائل المتاحة. ويُشير هذا الباحث أيضاً إلى أنه وعلى الرغم من أهمية تعرف أسباب النجاح والفشل في البرنامج فإن بالإمكان التوصية ببرنامج ناجح مع ترك أسباب

تحديد عوامل النجاح للمسؤول عن البرنامج وليس للمقوِّم. ولتلافي التركيز على الأهداف وإهمال التدقيق في النواتج والتأثيرات غير المقصودة التي أحدثها البرنامج، فقد رأى سكريفن ضرورة إجراء تقويم متحرر من الأهداف Goal Free Evaluation وفي هذا النوع يُركز المقوِّم على التعرف على الآثار التي أحدثها البرنامج والتي قد لا تكون ضمن أهدافه والعودة بعد ذلك إلى الأهداف الأصلية للمقارنة معها.

قائمة سكريفن للعملية التقويمية:

طوَّر سكريفن قائمةً تعكس رؤيته للعملية التقويمية تتضمن جوانب تشتمل عليها وأسئلة يجب الإجابة عنها (Scriven, 1983)، وفيما يلي مكونات هذه القائمة:

1- الوصف: أي يجب وصف ما يقوِّم، ففي تقويم البرنامج نذكر الأجزاء التي يتكون منها البرنامج وبيان العلاقة بين تلك الأجزاء ببعضها.

2- المستفيد: تحديد الجهة أو المسؤول الذي يطلب التقويم، ومن الذي إقترح التقويم وما الجهات ذات المصلحة من البرنامج.

3- المحتوى وخلفية البرنامج وأسلوب التقويم: أي فحص بيئة البرنامج، بيئة التقويم، نوع التقويم المطلوب وهل هو بنائي مستمر، نهائي أم مقارن أم غير ذلك.

4- المصادر: ما المصادر المتاحة للبرنامج وللتقويم سواء كانت مادية أو بشرية وما المتوفر والمستخدم منها؟

5- الوظيفة: ماذا يقدم البرنامج؟ ما وظيفته؟ ما الذي يُفترض أن يعمله، وما الذي يعمله فعلاً.

6- نظام تقديم الخدمات في البرنامج أو كيفية التشغيل: كيف يصل البرنامج إلى جمهوره؟ كيف يتم تطبيق البرنامج؟ ما الطرق المتبعة لضمان استمرارية خدمة البرنامج؟

7- المستهلك: من الذي يتلقى منافع البرنامج ويتأثر بهِ؟ من المتأثر به مباشرةً أو بصورة غير مباشرة؟
8- القيم والاحتياجات: ما القيم والاحتياجات لدى المتأثرين بالبرنامج؟ كيف يؤثر البرنامج في تلك القيم والاحتياجات ويتأثر بها؟
9- المعايير: ما المعايير المستخدمة للحكم على نجاح البرنامج؟ وهل هي موجودة فعلاً، أم يجب تطويرها؟
10- العمليات الداخلة: ما المحددات والتكلفة ذات التأثير على مستوى وفعالية تشغيل البرنامج.
11- النواتج: ما الآثار والنواتج المترتبة على البرنامج؟ مَن المتأثرين وما نوع التأثير الذي طرأ عليهم.
12- التعميم: هل يمكن تعميم البرنامج أو التوسع فيه؟ وما الظروف اللازمة لتعميمه؟
13- المقارنة مع الخيارات والبرامج البديلة: ما قيمة البرنامج مقارنةً بغيره، وما البدائل المتاحة؟ وهل يمثل البرنامج أفضلها؟
14- الدلالة: تركيب كل عناصر القائمة بأجمعها لتكوّن تصوراً معيناً عن دلالة البرنامج وأهميته العملية.
15- التوصيات: قد تكون مطلوبة أو غير مطلوبة في التقويم، وقد لا يتيسر تقديم توصيات رغم أنها مطلوبة.
16- التقرير: أي الالتزام بكل قواعد كتابة التقارير مثل الدقة، التوقيت المناسب، الصياغة الجيدة، اعتبارات خاصة بالتقويم وغيرها.
17- التقويم البعدي: أي تقويم التقويم Meta Evaluation، بهدف التعرف على مصداقية التقويم ويفضل أن يتم قبل التقرير النهائي، وقد يقوم به مقوِّم البرنامج نفسه أو مقوِّم آخر.

لاحظ سكريفن (1973) ومن خلال مشاركاته في تقويم برامج واسعة النطاق في الولايات المتحدة أن أخصائي التقويم ينشغل بدرجة كبيرة بالأهداف

المحددة مسبقاً والمعلنة في البرنامج دون إدراك إلى أنه ربما هناك نواتج غير متوقعة أو غير مطلوبة تبرز أثناء تنفيذ البرنامج لا تؤخذ بعين الاعتبار وعندئذ يصعب على المقوِّم الفصل بين الأهداف المعلنة والأهداف غير المتوقعة، ويوصي أن تهتم دراسات التقويم باكتشاف التأثيرات الفعلية للبرنامج، ومن هنا جاء اقتراحه لأسلوب (التقويم المتحرر من الأهداف) والذي يُريد فيه أن تظل أهداف البرنامج المراد تقويمه خافيةً - عن قَصد- على المقوِّم بحيث يركز اهتمامه على البحث عن التأثيرات المحتملة للبرنامج دون الاهتمام بأهداف البرنامج، ويضيف أن هذا النمط من التقويم أكثر موضوعيةً من النمط الموَّجه بالأهداف، ففي الأسلوب المقترح سوف لن يتأثر المقوِّم بتوجيهات المسؤولين عن البرنامج فيما يتعلق بالأهداف، بل ويرى أن من الضروري أن لا يكون المقوِّم هو أحد المسؤولين عن البرنامج لكي لا يتفاعل معهم ويتأثر بأفكارهم مما يؤدي إلى تحيزه في عملية التقويم، لكنه لا يمانع في أن يلتقي المقوِّم في المراحل الأولى من عمله بالمسؤول ليوضح له وصفاً موجزاً عن عمل البرنامج ليساعد بذلك في قيام المقوِّم بوضع خطته في ضوء هذه المعلومات الموجزة ويرسم صورة لخطة التقويم وكيفية حصوله على المعلومات المطلوبة ومقابلة المشاركين بالبرنامج من معلمين وطلبه، والقيام بتحليل الاختبارات المدرسية ونتائجها أو استخدام اختبارات متنوعة لتقويم النتائج (وهي النتائج التي تختلف عن تلك التي يسعى البرنامج لتحقيقها)، ومن ثم التعرف على التغيرات التي تحدث في اتجاهات الطلبة وتشخيص دوافعهم وسلوكيات أخرى يرى المقوِّم ضرورة قياسها والاستفادة منها ?وقد وجه بعض الباحثين النقد لأسلوب سكريفن في أنه قد يكون مفيداً في سياق التقويم الختامي Summa-tive Evaluation للمساعدة في الحكم على جودة منتج معين، وإجراء تحليل الكلفة وتحديد منفعة البرنامج، لكن سكريفن أكد على أن النموذج الذي يطرحه يمكن أن يلعب دوراً هاماً في قرارات تحسين البرنامج، إذ عندما توجد تأثيرات غير متوقعة أو جانبية في المراحل البنائية Formative لبرنامج معين فإنها تؤدي إلى إجراء التعديلات المناسبة على التأثيرات المرغوبة.

ب. نموذج (الهيئة الكلية) لستيك **Stake Countenance Model**

إنتقَدَ ستيك (Stake, 1967) نموذج تايلر المعتمد على الأهداف وأكد على ضرورة وجود اتصال مباشر بين المقوِّم والمستفيد من أجل اكتشاف الأمور المهمة في البرنامج المقوَّم وأخذها بالاعتبار عند إجراء التقويم. وقد انتقد أسلوب جمع المعلومات التقويمية من خلال التركيز على استخدام الاختبارات والأدوات المقننة والتجارب، لكنه أشار إلى أنه يمكن حصر استخدامها في أضيق مجال من مجالات البرنامج.

لقد وضع ستيك هذا النموذج الخاص بالتقويم عام 1967 وأعقبه بوضع نموذج آخر عام 1975 أسماه نموذج (التقويم المتجاوب أو المستجيب Respon- sive Evaluation) وأكد على ضرورة استخدام أساليب شبيهة "بالإثنوغرافية Ethnography" لجمع المعلومات أو استخدام أسلوب "دراسة الحالة Case Study" ويمكن من خلال عملية التقويم أن يجمع المقوِّم آراء الآخرين وأحكامهم عن البرنامج مضيفاً إليها رؤيته وأحكامه ليقدمها للمستفيد بأسلوب مناسب. أما نموذج (الهيئة الكلية) فقد قدمه كإطار للتقويم يمثل المنظور أو الصورة التي يجب أن يكون عليها التقويم ويتكون من العناصر الآتية:

1- وصف البرنامج وتقويمه.
2- تحديد مصادر المعلومات المختلفة.
3- تحليل التناسق والتلازم في عناصر البرنامج.
4- تحديد المعايير ذات العلاقة.
5- الاستخدامات المختلفة لنتائج التقويم.

ويلاحظ أن ستيك ركز في هذا النموذج على عملين أساسيين هما: الوصف Description والحكم Judgment ولذلك يسميه أحياناً بنموذج ("الملامح العامة للتقويم"). ويضيف ستيك أنَّ العمليتين السابقتين تمران بثلاثة محاور (مراحل) من حياة البرنامج التعليمي وهي:

السوابق Antecedents، والتفاعلات Transactions، والنواتج Outcomes.

ويُقصد بالسوابق الظروف الموجودة قبل بدء عمل البرنامج والتي قد يكون لها تأثيراً على البرنامج حيث يتعين بالمقوِّم جمع معلومات عن تلك الظروف لتوفير خلفية وافية عن البرنامج واستخدامها في تخطيط أو تنفيذ عملية التقويم أو تفسير بعض الظواهر التي يبني عليها المقوِّم بعض الاستنتاجات الضرورية عند كتابة التقرير النهائي للتقويم.

ويقصد بالتفاعلات الأمور التي تحدث أثناء عمل البرنامج ولها أثر كبير على سير البرنامج ومنها العلاقات السائدة بين الأفراد العاملين في البرنامج وتوزيع الأعمال بينهم وأساليب التعامل بين المسؤول وبينهم، ويستطيع المقوِّم ملاحظة تلك التفاعلات التي تفيد في تفسير جوانب القصور أو النجاح أثناء تنفيذ البرنامج. أما النواتج فيقصد بها الآثار التي أحدثها البرنامج، أي ما ينتج عنه وينعكس على سلوك واتجاهات المستفيدين من البرنامج أو تحصيلهم الدراسي إن كانوا طلاباً، ولتحديد تلك التأثيرات يلجأ الباحث إلى استخدام أساليب قياس مناسبة أو الشعور بها من خلال ملاحظات موضوعية.

ويُشير ستيك فيما يخص عملية الوصف Description أنه يجب أن يكون شاملاً لكل مراحل البرنامج وللمحاور الثلاثة سابقة الذكر، وأن يفصل المقوِّم بين ما هو مخطط له وما هو متحقق في تلك المحاور، ثم يجري تحليلاً للمعلومات الوصفية للوقوف على مدى التوافق بين المستهدف والملاحظ من مهام ووظائف البرنامج. أما فيما يخص عملية إصدار الأحكام Judgments فيحمل المقوِّم والمسؤولين عن البرنامج مهمة الاتفاق على تحديد المعايير التي سيتم استخدامها في التقويم لأجل إصدار أحكام سير البرنامج في ضوئها. وهذه المعايير قد تكون (نسبية) مثل خصائص المراحل التي يمر بها البرنامج (السوابق، التفاعلات، النواتج) مقارنةً ببرامج مماثلة وقد تكون (المعايير) مطلقة تعتمد على توقعات المسؤولين حول البرنامج وما يرونه في البرنامج من وجهة نظرهم الخاصة. ويمكن تلخيص عمل المقوِّم وفقاً لهذا النموذج بالآتي: -

أ. القيام بجمع معلومات وصفية عن البرنامج وتحليلها بشكل دقيق وموضوعي.
ب. تحديد المعايير المطلقة والنسبية التي يجري في ضوئها إصدار الأحكام.
جـ. تقدير مدى الإيفاء بتلك المعايير.
د. إصدار الأحكام المتعلقة بالبرنامج سواء من قبل المقوِّم وحده أو بالتعاون مع المسؤولين عن البرنامج أو بمشاركة أفراد آخرين متأثرين بذلك البرنامج.

جـ. نموذج التقويم المتجاوب Responsive Evaluation

وضع هذا النموذج من قبل ستيك (Stake) ويعتمد في عملية جمع المعلومات عن البرنامج المقوَّم على النموذج السابق الذي وَضَعَهُ وهو نموذج الهيئة الكلية Countenance Model، ويركز على تحديد جوانب القوة والضعف في البرنامج المقوَّم وليس فقط التأكد من مدى تحقق أهداف البرنامج. ونظراً لأن هذا النموذج يتمحور حول المستفيد من التقويم فإن عملية التقويم لا تقف عند حد الأهداف ومدى تحقيقها كما أسلفنا، بل تمتد إلى الأحداث والقضايا المهمة في البرنامج وجمع معلومات عنها من خلال إجراءات تعتمد على الملاحظة المباشرة والاستماع إلى آراء المستفيدين وعلاقتهم بالبرنامج وهم في الظروف الطبيعية Naturalistic أي غير المصطنعة تماماً كما يفعل الباحث الأنثروبولوجي في التعامل مع الأفراد الذين يركز على دراسة ثقافاتهم وعلاقاتهم الاجتماعية وعاداتهم وتقاليدهم، ولذلك يبقى الاتصال بين المقوِّم والمستفيد مستمراً وغير رسمي ومستجيباً لمتطلبات الجانبين خلال مراحل التقويم. وعندما يضع المقوِّم أحكامه عن البرنامج تكون تلك الأحكام مستندة إلى آراء المستفيدين أو المتأثرين بالبرنامج أكثر من استنادها إلى آراء المقوِّم نفسه. ويرى ستيك الذي قدَّم هذا النموذج في مؤتمر علمي عن (مداخل التقويم الاستجابي) الذي عُقدَ في السويد عام 1974 أن إجراء هذا النوع من التقويم يتطلب تطوير خطة للملاحظات والتفاوضات يشارك فيها أكثر من ملاحظ ثم يقوم المختص في التقويم بوضع (مقترح وصفي) و(تصور للنتائج)

و(رسوم وأشكال بيانية) وغير ذلك ويحاول من خلال ذلك التوصل إلى ما يهم المعنيين وجمع انطباعات وتصورات الأفراد حول ما يودونه من البرنامج ويعرضها على العاملين بالبرنامج ليتعرف على آرائهم عن تلك المعلومات وتقييم مدى مواءمتها، ويقوم بتعديل المعلومات استناداً إلى هذه الآراء والتقييمات ويكرر ذلك إلى أن يتحقق من اكتمالها والإتفاق عليها. ويمكن إجراء ذلك بوسائل تواصل عادية، ثم يُعد تقريراً نهائياً يتناول تخطيطاً مبدئياً وتركيزاً متتابعاً للدراسة التقويمية وبذلك يكون لدى المقوِّم منظوراً شاملاً للبرنامج، ورغم أن هذه الخطوات تسير بشكل سريع نوعا ما فإنها قد تحقق الغرض من هذا النوع من التقويم الذي يتميز بأنه سريع الاستجابة وهذه الاستجابة تكون أكبر لمتطلبات الجهة التي تطلب معلومات معينة عن البرنامج.

ويتميز هذا النموذج بأنه يعتمد اعتماداً كبيراً على الوصف اللفظي للأحداث الجارية في البرنامج، ويقدم معلومات يستند إليها صانع القرار في أحكامه المتعلقة بالبرنامج، ويحاول بشتى الطرق أن يعكس البرنامج صورة واضحة مستندة على أكثر من منظور قد تعطي صورة مختلفة عما يكشفه الفحص التفصيلي لبعض مكونات البرنامج. إلا أن ما يؤخذ على هذا النموذج من بعض خبراء التقويم أنه يركز على البيانات الذاتية التي تصف ما يحدث في البرنامج ومدى جودته استناداً إلى رؤية الأطراف المعنية، وبذلك لا يكون هناك فصل بين قيمة البرنامج ورؤية تلك الأطراف عن البرنامج، يضاف إلى ذلك أن المقوّم يعتمد على خبرته عند تدوين الملاحظات التي يستخدمها وكذلك في تفسيرها.

وأخيراً واستناداً إلى ما تقدم يمكن تلخيص الأسس التي يقوم عليها هذا النموذج بالشكل الآتي:

1- اهتمامه بالقضايا أو المسائل غير المؤكدة، والمسببة للقلق فيما يخص برنامج معين مما تدفع إلى إجراء عملية تقويمه.

2- يعتمد في جمع المعلومات عن البرنامج المقوِّم على الخطوات والأسس التي وردت في نموذج (الهيئة الكلية).

3- استفادة المقوِّم من الأفراد المستفيدين كمراقبين للبرنامج وجمع المعلومات عن البرنامج من خلال ملاحظاتهم وانطباعاتهم بصفتهم خير مصدر للمعلومات وأفضل أداة لجمع المعلومات.

4- إمكانية الثقة بالمعلومات التي تجمع عن البرنامج لأنها ترد من مصادر متنوعة دون اعتمادها على مصدرٍ واحد.

ثالثاً: النماذج المساعدة على اتخاذ القرارات

تخدم هذه النماذج متخذي القرارات من إداريين وصانعي السياسات التربوية ومتخذي القرارات وحتى المعلمين الذين هم بحاجة ماسة إلى معلومات حول الناتج التربوي أو المنهج. والمبدأ الذي ترتكز عليه هذه النماذج هو أن المعلومات التقويمية ولاسيما (النواتج) جزءً لا يتجزأ من عملية صنع القرار وتوفيرها للمسؤولين بما يُساعد على إحداث تغيير إيجابي في البرامج التربوية. ويلاحظ من خلال استقراء مرتكزات هذه النماذج أن المقوِّم لا يشغل نفسه بالحكم على جدارة البرنامج بل أن دوره ينحصر في تقديم المعلومات لصاحب القرار الذي في ضوئها سيحكم على جدارة البرنامج. ومن أشهر الباحثين الذين وضعوا أسس هذه النماذج كل من آلكن وستيفلبيم (Alkin وStafflebeam) وسنقدم في الصفحات القادمة توضيحاً لنموذج كل منهما على انفراد.

أ. نموذج مركز دراسات التقويم في جامعة كاليفورنيا - لوس أنجلس

(UCLA Center for the Study of Evaluation)

أشرف الكن Alkin على تصميم هذا النموذج في المركز الذي كان يعمل فيه بجامعة كاليفورنيا - لوس أنجلس عام 1966 ويتضمن تحديد الإجراءات التي يُفترض أن تتخذ في كل مرحلة من مراحل تطور البرنامج ثم اختيار المعلومات اللازمة وتوفيرها لمتخذ القرار لمساعدته في اتخاذ القرارات والإجراءات المناسبة لعمل البرنامج وفقاً للآتي:

1- مرحلة تقدير الاحتياجات: Needs Assessment

ويركز المقوِّم في هذه المرحلة على معرفة مدى وفاء البرنامج بالاحتياجات الخاصة للمستفيدين منه والمجتمع المحلي أيضاً وذلك من خلال إجراء مقارنة بين أهداف البرنامج والنواتج الخاصة به، ثم يحدد الفرق بينهما لتحديد الاحتياجات التي تعتبر من ناحية أخرى أهداف يُفترض العمل على تحقيقها، وبذلك تكون الأهداف منبثقة من الحاجات مما يسهل تنفيذها بدقة.

2- مرحلة التخطيط: Program Planning

وفي هذه المرحلة يحاول المقوِّم توفير معلومات عن نوعية البرامج التي تلبي الاحتياجات التي حددت في المرحلة السابقة سواء تحديد برامج تدريبية لرفع مستوى كفاءة العاملين في البرنامج، أو وضع برامج تدريبية أكثر فاعلية لاستخدامها من قبل المعلمين. كما يقوم المقوِّم بتزويد صانع القرار بمعلومات تساعده في الاختيار بين البدائل المتعددة للتطوير وعليه اختيار الأنسب للبرنامج في تحقيق الاحتياجات. وتتضمن هذه المرحلة تحديد المدة الزمنية اللازمة لاستكمال تنفيذ البرنامج، وكلفة تحقيق أهدافه، وتحديد مستلزمات تنفيذه بشرية أو مادية.

3- مرحلة تقويم العمليات: Process Evaluation

بعد أن تم تحديد الاحتياجات ووضع خطة البرنامج تبدأ عملية تنفيذ البرنامج وهنا يحاول المقوِّم تحديد مدى ملاءمة العملية للمواصفات التي تم تحديدها في الخطة، لذلك يُعتبر تقويم العمليات بمثابة تقويم بنائي Formative Evaluation تهدف للكشف عن نواحي القصور والقوة التي تبرز خلال التنفيذ وكأنه عملية مراقبة مستمرة Monitoring عن طريق جمع بيانات، ومتابعة التفاعلات المتبادلة بين عناصر البرنامج، ومراجعة الجداول الزمنية لتنفيذ البرنامج وتحديد فاعلية الأدوات والأجهزة المستخدمة وأثرها في زيادة كفاءة وخبرة الأفراد، وكفاية المشرفين على سير العمل. أي التأكد مما إذا كان البرنامج يسير في الاتجاه المحدد له في تحقيق الأهداف أم لا.

4- مرحلة تقويم النتائج: **Outcome Evaluation**

ويجري في هذه المرحلة تجميع المعلومات المتعلقة بمدى تحقيق نتائج البرنامج للأهداف التي حددت سلفاً وربط ذلك بتقدير الاحتياجات، والتخطيط وتقويم العمليات التي أُجريت في المراحل السابقة من أجل مساعدة متخذ القرار في الحكم على نجاح أو تلكؤ البرنامج في تحقيق أهدافه وما ينشأ عن ذلك من قرارات تتعلق بتعديله أو إيقافه أو التوسع فيه. ولا شك أن ما سبق يُعتبر من خصائص التقويم الختامي Summative Evaluation.

وأخيراً لابد من الذكر أن هذا النموذج نال اهتماماً بالغاً وتم استخدامه بشكل واسع من قبل المقوِّمين في الولايات المتحدة. كما أن هذا النموذج يتفق مع نموذج (ستفلبيم) في أوجه كثيرة منها تحديد مراحل التقويم، وحرصه على مساعدة متخذ القرار، رغم أنه أجرى نوعاً من التعديل في مجال (تقويم العمليات) بحيث أصبح يشمل تقويم التنفيذ Implementation Evaluation، وتقويم التقدم Progress Evaluation وتقويم النواتج الوسطية التي لم يؤكد عليها نموذج ستفلبيم، ويرى البعض أن نموذج الكن Alkin يركز بدرجة أساسية على التساؤل: "متى يجرى التقويم؟". بينما تركز معظم النماذج الأخرى على التساؤل: "كيف يُجرى التقويم؟".

ب - نموذج ستفلبيم Stufflebeam Model (CIPP)

وضع هذا النموذج الباحث الأمريكي ستفلبيم (1969) مستنداً إلى خبرته الواسعة في تقويم البرامج التربوية وتحليلاته الناقدة لنماذج التقويم المتعددة السابقة له، والذي يرى أن التقويم هو: العملية التي نُحدد بها المعلومات الوصفية والتحكيمية ونجمعها ونُوفرها لإصدار الحكم على قيمة الموضوع المقوَّم وجدارته من أجل توجيه القرار لأغراض المحاسبية وزيادة فهمنا للظاهرة المدروسة (Stufflebeam, 1983).

ويلاحظ من خلال استقراء التعريف أنه يخدم ثلاثة أغراض هي: توجيه القرار، وتوفير معلومات تخدم مبدأ المساءلة: Accountability إضافةً إلى فهم

الظاهرة قيد التقويم. والتقويم هنا ليست حادثة عرضية بل عملية منظمة وهادفة تشمل ثلاث خطوات متداخلة ومتكاملة مع بعضها وهي: تحديد المعلومات الوصفية والتحكيمية وتوثيقها وتقديمها لخدمة الأغراض المختلفة لصانع القرار، ويشمل التقويم أهداف البرنامج، وتصميمه وتنفيذه والأثر المترتب عليه أو الناشئ عنه وفقاً لمعيارين هما الجدارة والقيمة. وتعني الجدارة تميز البرنامج من حيث كفاءته وفاعليته، أما القيمة فتعني مدى الفائدة المتحققة من البرنامج وفقاً لتوقعات المسؤولين عنه والحاجة إلى الخدمة التي يقدمها، والجدوى من البرنامج. وبالنظر لأن الهدف الأساسي لوضع هذا النموذج هو مساعدة المسؤولين في اتخاذ القرارات المرضية حول البرنامج التربوي، فإن واضعه صنف تلك القرارات في أربعة أصناف هي:

(1) قرارات التخطيط Planning decisions الخاصة بتحديد وتطوير الأهداف الرئيسية والإجرائية للبرنامج التربوي.

(2) قرارات البناء Structuring decisions الخاصة بتركيبة البرنامج وإجراءاته.

(3) قرارات التنفيذ Implementation decisions وتتعلق بقضايا استخدام وضبط وإعادة فحص وآليات التنفيذ.

(4) قرارات التَدوير Recycling decisions وتتعلق بالحكم على البرنامج من خلال نتائجه، واتخاذ القرار بشأنه سواء باعتماده، أو إجراء تغييرات فيه أو إيقافه.

وقد اقترح (ستفلبيم) أربعة أنواع من التقويم تتناغم مع أنواع القرارات الأربعة السابقة الذكر هي:

تقويم السياق (Context Evaluation) وتقويم المدخلات (Input evalua-tion) وتقويم العمليات أو الإجراءات (Process Evaluation) وتقويم المخرجات (Product Evaluation) وبذلك يعرف هذا المدخل في التقويم بـ(CIPP). وسوف يتم إعطاء شرح موجز لطبيعة هذه القرارات:

تقويم السياق: يهدف إلى تقديم معلومات وبيانات تساعد في اتخاذ

القرار في مرحلة التخطيط من خلال تحديد الاحتياجات التي يجب أن يلبيها البرنامج والتي تساعد في صياغة أهدافه.

تقويم المدخلات: يهدف إلى تقديم معلومات وبيانات تخدم متخذ القرار في مرحلة البناء (Structuring Decisions) من خلال تحديد المصادر المتوفرة، والاستراتيجيات المختلفة لتنفيذ البرنامج.

تقويم العمليات: يهدف إلى تقديم معلومات وبيانات تساعد في اتخاذ القرارات في مرحلة التنفيذ. من خلال تقديم معلومات عن سير خطة التنفيذ والعوائق التي تعترضها، وكيف يمكن مراجعة خطة التنفيذ في ضوء ذلك ومن ثم تعديل طريقة تنفيذ البرنامج.

تقويم النواتج أو المخرجات: يهدف إلى تقديم معلومات عن جودة المخرجات والتي تساعد على اتخاذ القرارات حول فاعلية وكفاءة مخرجات البرنامج.

وتتكون عملية التقويم عند ستفلبيم من الخطوات الآتية:

أ- **وضع محددات التقويم (Focusing the Evaluation)** ويتضمن:

1- تحديد المستويات المختلفة من متخذي القرار الذين سوف يستخدمون نتائج التقويم (مدراء مدارس، موجهين، مدراء عموم).

2- وضع تصور للقرارات التي سوف تتخذ ودرجة أهميتها والبدائل المتوفرة لكل مستوى من مستويات متخذي القرار.

3- تحديد معايير لكل قرار من خلال بيان المتغيرات التي سوف يتم قياسها والمعايير التي سوف تستخدم للمفاضلة بين البدائل.

ب- **جمع البيانات وتتضمن**:

1- تحديد مصادر البيانات التي سوف يتم جمعها.

2- تحديد أدوات جمع البيانات وآليات جمعها.

3- تحديد إجراءات اختيار العينة المستهدفة.

4- وضع جدول زمني لجمع هذه البيانات.

جـ- تنظيم البيانات ويتضمن:

1- وضع إطار للبيانات التي سوف يتم جمعها وتبويبها.

2- تحديد وسائل تحليل البيانات.

د- تحليل البيانات وتتضمن:

1- اختيار أساليب التحليل المناسبة.

2- تحديد طرائق إجراء التحليلات.

هـ- كتابة التقرير ويتضمن:

1- تحديد الجهة التي يستهدفها التقرير أو الجمهور المستهدف.

2- تحديد وسائل وأساليب تزويد المستهدفين أو الجمهور بالمعلومات.

3- تحديد شكل التقرير في صورته النهائية.

4- وضع جدول زمني لكتابة التقرير.

و- إدارة عملية التقويم وتتضمن:

1- تحديد الموارد البشرية والمادية اللازمة لعملية التقويم.

2- تقويم عملية التقويم ذاتها من حيث كفاءة التصميم التجريبي المستخدم في الحصول على معلومات صادقة وثابتة، وموضوعية وفي الوقت المناسب.

3- تحديد آليات تحديث بيانات التقويم مرحليا.

4- تحديد الميزانية المطلوبة لعملية التقويم (Stufflebeam, 1983)

يتضح مما سبق إن أهم ما يميز هذا النموذج تركيزه على جمع المعلومات ذات العلاقة بأسئلة الدراسة من خلال تواصله الدائم مع متخذي القرار. فعملية التقويم تتطلب جمع أكبر قدر من المعلومات والبيانات التي قد تفيد في تعويض ما قد يتبين فيما بعد أن بعض هذه المعلومات قد لا يفيد موضوع التقويم. كما إن التقويم يجب أن يأخذ مكانه في أي مرحلة من مراحل تطوير البرنامج وليس بالضرورة الانتظار إلى أن يتم تطبيقه أو الانتهاء منه. فالعائد من التقويم بعد انتهاء البرنامج يظل محدودا وإن كان يخدم أغراضا

مختلفة. وهذا النموذج هو المدخل المفضل لدى التربويين وصناع القرار لأنه يركز على توفير المعلومات الضرورية لمتخذي القرار. ويخدم هذا النموذج أيضا عملية تقويم كل خطوات بناء المنهج ويصاحب خطوات بنائه وتطوره وبعد ذلك تطبيقه وفي كل خطوة من الخطوات يقدم معلومات مهمة أولا بأول لمتخذي القرار لتصحيح مسار البرنامج أو المنهج.

ولعل أهم عيوب هذا النموذج عجز القائم بعملية التقويم أحيانا عن الاستجابة لبعض القضايا المهمة في عملية التقويم لأنها لا تتماشى مع اهتمامات متخذي القرار وهم الذين يتحكمون في عملية التقويم وفقاً لهذا النموذج، وخاصة إذا لم يتم تحديد الأولويات مما يؤدي إلى هدر كبير في الأموال وبدلا من أن تصبح عملية التقويم جزءا من بناء المنهج وتعمل على تحسينه وتجويده فإنها قد تستهلك جزءا كبيرا من ميزانيته. وأخيرا فإن هذا النموذج يعتمد على أن هذه القرارات يمكن تحديدها مسبقا، وأنها تبقى ثابتة وغير متغيرة أثناء سير عملية التقويم وحتى الانتهاء منها.

ج - نموذج التقويم المرتكز على الخبراء (Experts-oriented Model)

يعد النموذج المرتكز على الخبراء في التقويم من أقدم مداخل التقويم وأكثرها شيوعا ويعتمد كليا على رأي الخبراء حول المؤسسات التعليمية والبرامج والمخرجات والأنشطة التعليمية. فعلى سبيل المثال يتم تقييم جودة منهج معين بواسطة خبراء في المادة التعليمية وذلك من خلال ملاحظة تطبيق المنهج وتحليل محتواه من أجل التوصل إلى حكم عن جدواه وكفاءته. فإذا كان الخبراء يسهمون بشكل أو بآخر في كل نماذج التقويم التي ذكرناها إلى الآن إلا أنهم يمثلون حجر الأساس في هذا المدخل، فهم يُعتَبَرون المصدر الرئيس لكافة البيانات والمعلومات التي تعتمد عليها عملية التقويم.

وتوجد عدة أمثلة لهذا النموذج، إلا أننا سنتعرض لنموذج واحد فقط في هذا الفصل وهو أشهرها على الإطلاق ألا وهو نموذج الاعتماد

الأكاديمي academic accreditation والاعتماد الأكاديمي يمكن تعريفه بأنه العملية التي يتم من خلالها الاعتراف بمؤسسة غالبا ما تكون حديثة العهد أو صغيرة من قبل أخرى كبيرة أو مجلس اعتماد أو رابطة معينة. ويمثل مجلس الاعتماد بسلطنة عُمان الذي صدر به المرسوم السلطاني رقم (2001/74) الجهة التي تمنح الاعتماد الأكاديمي المؤسسي والبرامجي لمختلف مؤسسات التعليم العالي بالسلطنة حكومية كانت أم خاصة. ويهدف الاعتماد الأكاديمي إلى التحقق من أن المؤسسات التعليمية المختلفة التي يتم اعتمادها تطبق معايير الجودة التي تضعها جهة الاعتماد. والاعتماد ليس مقتصرا على المؤسسات التعليمية فحسب بل إن هناك مهن أخرى مهتمة بقضية الجودة كالطب والهندسة والقانون. وغالبا ما يكون الاعتماد الأكاديمي اختياريا نابعا من حرص المؤسسة على جودة خدماتها وبالتالي جودة مخرجاتها (الصارمي، 2003).

ويستند الاعتماد على الفرضية القائلة بأن الأكفأ والأقدر على الحكم على أنشطة أصحاب مهنة معينة وجودة عملهم هم المتنسبون إلى تلك المهنة أي من يعمل بتلك المهنة. وهؤلاء يقتصر عملهم على الزيارات الميدانية للمؤسسات المختلفة ووضع المعايير والمواصفات التي يعتمد عليها للحكم على تلك المؤسسات (Worthen & Sanders, 1984) وفي الماضي ركز اعتماد المؤسسات التعليمية على ثلاثة محاور رئيسة:

1- توفير الخدمات المناسبة.

2- المؤهلات الأكاديمية والمهنية للهيئة الأكاديمية بالمؤسسة.

3- ملاءمة التصميم التعليمي والبيئة التعليمية المستخدمة في المؤسسة.

ولم تحظ المخرجات أو النواتج التعليمية في عملية الاعتماد في ذلك الوقت بالاهتمام. ولكن الانتقادات الموجهة لمؤسسات الاعتماد لإهمالها هذا الجانب المهم من العملية التعليمية وأدت إلى تبني هذه المؤسسات للمخرجات كمحور رابع في عملية الاعتماد، وكان ذلك بداية عام 1972 (Dickey & Miller, 1972) وقد لخص سكريفن (Scriven, 1984) الخطوات التي تتم فيها عملية الاعتماد في الآتي:

1- إيجاد معايير الاعتماد بواسطة الجهة المسؤولة عن الاعتماد وإعلانها للمؤسسات الراغبة في الاعتماد.
2- قيام المؤسسة الراغبة في الاعتماد بدراسة ذاتية (Self Study)
3- قيام فريق من الخبراء من مؤسسة الاعتماد بزيارات ميدانية للمؤسسة الراغبة في الاعتماد.
4- إعداد تقرير عن المؤسسة شاملا التوصيات من قبل فريق الخبراء.
5- إخضاع التقرير المعد عن المؤسسة بواسطة فريق الخبراء إلى التحكيم بواسطة محكمين محايدين peer review وذلك للتأكد من سلامة المنهجية المتبعة في الدراسة وجودة أدوات جمع البيانات.
6- كتابة التقرير النهائي لمؤسسة الاعتماد متضمنا قرارها باعتماد المؤسسة أو عدم الاعتماد مع إعطاء المبررات لذلك.

وبالرغم من أن توصيات مؤسسات الاعتماد غير ملزمة للمؤسسات الراغبة في الاعتماد إلا أن عملية الاعتماد ساهمت في إعادة هيكلة المدارس الثانوية في بعض الدول كالولايات المتحدة الأمريكية (Worthen & Sanders, 1987)

ويوجد هناك نوعان من الاعتماد:

1- **الاعتماد المؤسسي**: ويعني اعتماد المؤسسة ككل بما تحويه من برامج وأنشطة، فالاعتماد هنا يعني بأن مؤسسة أو هيئة الاعتماد قد أقرت بأن المؤسسة التعليمية بشكل عام تراعي وتطبق المعايير التي وضعتها هيئة الاعتماد.
2- **الاعتماد البرامجي أو المتخصص**: ويعنى باعتماد البرامج المتخصصة داخل المؤسسة، فيمكن أن يوجد داخل المؤسسة برنامج أو أكثر معتمد ولكن ليس بالضرورة أن تكون المؤسسة الأم معتمدة، بمعنى أن الاعتماد المؤسسي ليس شرطاً ضرورياً للاعتماد البرامجي.

ولعل أهم اسهامات هذا النموذج وخاصة المتعلق منه بالاعتماد الأكاديمي هو إشاعة ثقافة الجودة في مؤسسات التعليم العالي من خلال تبني معايير وإرشادات لتقييم فعالية وكفاءة تلك المؤسسات ودفعها إلى القيام بدراسات

ذاتية self-study وتقويمية. ولعل أهم من ذلك كله طمأنة الطلاب وأولياء الأمور وأصحاب الأعمال بأن المؤسسات المعتمدة لها أهداف واضحة ومحددة، ولديها بيئة تعليمية تساعد على تحقيق تلك الأهداف، وأن الأهداف الموضوعة يتم تحقيقها بصورة مرضية، ولديها العدد الكافي والمناسب من الهيئات الأكاديمية والأكاديمية المساندة والموظفين، وأنها مؤهلة في الاستمرار في هذا المجال. مما ينتج عنه تنامي الثقة في تلك المؤسسات لدى العامة وأصحاب الأعمال. كما تعمل مؤسسات وهيئات الاعتماد كهيئات استشارية مساعدة للمؤسسات قيد الإنشاء أو المؤسسات الحديثة. كما إن وجود مؤسسات اعتماد في دول العالم الثالث من شأنه التقليل من الإشراف المباشر من قبل مؤسسات الدولة على مؤسسات التعليم المختلفة وما ينتج عن ذلك من بيروقراطية وعدم ممارسة مؤسسات التعليم حرياتها الأكاديمية بالشكل الذي يضمن لها الاستقلالية التامة.

ولعل أهم عيوب عملية اعتماد المؤسسات التعليمية هو ارتفاع تكلفتها نظرا للزيارات المتعددة التي يتوجب على فريق الاعتماد القيام بها للمؤسسة قيد الاعتماد والإجراءات المصاحبة لها. كما يعتمد فريق الاعتماد كثيرا على الدراسة الذاتية الأولية التي عادة ما تعدها المؤسسة عن نفسها والتي عادة ما تكون متحيزة ولا تعكس واقع المؤسسة. كما أشار سكريفن (Scriven, 1984) إلى ثلاثة عيوب تواجه عملية الاعتماد: (1) خلو معايير الاعتماد من أوزان نسبية حيث إن جميع المعايير تعامل بنفس الأهمية، (2) التحيز الإداري المصاحب لعملية تشكيل فريق التقويم أو المراجعة، (3) الإجراءات التي تعوق إسهامات المنتقدين للبرنامج في عملية التقويم. كما أن زيارات فريق الاعتماد للمؤسسة قيد الاعتماد عادة ما تكون قصيرة وشكلية ولا تفضي إلى معلومات قيمة عن المؤسسة.

رابعاً: النماذج النوعية في التقويم Qualitative Model in Evaluation

سادت الأساليب الكمية في البحث والتقويم الميدان التربوي منذ بدايات القرن الماضي وتطورت بشكل متسارع في منتصف القرن ذاته لاسيما بعد شيوع المنهج التجريبي وشبه التجريبي في مجال علم النفس الذي تطور بشكل كبير بعد إصدار كامبل وستانلي (Campbell & Stanley, 1966) لكتابهما الخاص بالتصميم التجريبي وشبه التجريبي على الرغم من تأكيد العديد من الباحثين التربويين بأن استخدام المنهج التجريبي ليس ممكناً دائماً، إلا أن وجاهة هذا المنهج والصورة الحسنة له جعلت الكثيرين من المهتمين بشؤون التقويم التربوي يعتمدون على الأساليب الكمية اعتماداً كبيراً.

ولكن مع سبعينيات القرن الماضي بدأ الاهتمام بالتقويم النوعي والاعتماد عليه يتزايد وتوقع كل من بودجان وبلكان (Bodgan & Bilkan, 1982) أن استخدام التقويم النوعي في المجال التربوي سوف يتزايد وسيصبح له شأن عظيم في القريب العاجل، وهذا بالفعل هو الذي حصل على الواقع حيث تزايد استخدام الأساليب النوعية في أغلب المجالات الإنسانية كالاجتماع والتربية وعلم النفس والإدارة وغيرها وأصبح الباحثون والمقوّمون يعتمدون على المواقف الطبيعية كاستخدام المنهج الأثنوغرافي Ethnography أي المنهج المتعلق بالوصف الميداني الواقعي، والذي اهتم به علماء الأنثروبولوجيا الثقافية الاجتماعية، المكون من مجموعة عناصر أساسية لخصها كل من (Cook & Reichardt, 1979) بالآتي:

أ. أسلوب استكشافي مفتوح عن مشكلة البحث.

ب. مشاركة مكثفة للباحث في المحيط الاجتماعي المطلوب دراسته كملاحظ ومشارك بأساليب متنوعة.

جـ. استخدام أساليب بحث مكثفة مع التأكيد على الملاحظة بالمشاركة والمقابلات المعمقة.

د. محاولات نقدية لفهم الحوادث والأفكار التي يحملها الأفراد الواقعين تحت الملاحظة.

هـ. إطار تفسيري يؤكد على الدور المهم للبيئة التي تقرر السلوك الكلي الذي يحكم العلاقات المتداخلة للسلوكيات والحوادث المكونة للنظام الوظيفي السائد في المجتمع المدروس.

و. تقديم نتائج البحث مكتوب بشكل وصفي طبيعي يفسر الأحداث ويصف الواقع بحيث يشعر القارئ وكأنه يعيش الواقع نفسه.

ومن المناهج الأخرى التي يعتمد عليها الباحث أو المقوِّم النوعي دراسات الحالة Case Study التي تهتم بتسليط الضوء المكثف على الحالة الواقعة تحت التقويم وجمع معلومات مفصلة عنها سواء من سجلات أو ملاحظات أو أشخاص لهم علاقة وثيقة بالحالة واستخدام أدوات البحث التي تعتمد على الملاحظات والمقابلات الفردية أو الجماعية حيث يتمكن الباحث من جمع معلومات أكثر تنوعاً ودقةً حول الموضوع مما قد توفره الأساليب الكمية من خلال الاستبيانات والمقاييس والاختبارات المقننه التي قد لا تكشف عن حقيقة الظاهرة بقدر ما تكشف عن استجابات قد لا تكون حقيقية، بل قد تكون مجرد إشارات يضعها المستجيب في المكان المخصص لها على تلك الأدوات، وفي كثير من الأحيان تعكس آلية الاستجابة وليس حقيقتها. لقد أصبح دور المقوّم النوعي مشاركاً وليس ضابطاً أو متحكماً في عملية التقويم، ويقوم بدور المعلم والمتعلم ليكتشف الواقع على حقيقته ويقدم الصورة الواضحة المتكاملة له، وأقرب مثال على ذلك ما يقدمه الباحث الأنثروبولوجي من حقائق واسعة تكشف عن تفاصيل الثقافات والعلاقات الاجتماعية في المجتمعات التي يبحث فيها بدرجة تفوق أي أسلوب آخر يسعى للوصول إلى الهدف ذاته.

إن منهجية التقويم النوعي تعتمد على السياق العام الذي يجري فيه التقويم، ويكون المقوِّم مسؤولاً عن أدواره التي يجب أن تتحلى بالجوانب الأخلاقية والمعنوية التي تشترط الموضوعية في نقل الحقائق الملاحظة رغم اعتمادها على التأويل والتفسير لآراء الأفراد من ذوي العلاقة بالبرنامج المراد تقويمه وجمع المعلومات منهم بشكل مباشر ووجهاً لوجه، والرجوع إلى

السجلات الرسمية المتعلقة بالبرنامج والقائمين عليه، بما يوفر للمقوِّم معلومات كبيرة في حجمها ونوعها وما يمكن أن يؤدي ذلك في تقديم صورة واضحة وتفصيلية عن البرنامج المقوَّم.

وأخيراً لابد من الاشارة الى أنه كما أن هناك نماذج عديدة للتقويم الكمي ظهرت في المقابل أنواع أخرى من نماذج التقويم النوعي نذكر منها على سبيل المثال لا الحصر النماذج الآتية:

أ- التقويم التنويري Illuminative Evaluation

تطور هذا النموذج على أيدي كل من بارلت (Parlett, 1975) وهاميلتون (Hamilton, 1977) حيث يتميز بتقديم منهجية مرنة تراعي الموارد والفرص المتاحة وتستخدم أساليب متنوعة تعتمد على الأنثروبولوجيا الاجتماعية التي لا تهتم بقياس النواتج التربوية بقدر اهتمامها بالتفاصيل المتعلقة بأهمية وجود البرنامج وعملياته الأساسية والإنجازات التي يقدمها للأفراد المستفيدين منه والصعوبات التي يواجهها، أي أن هذا الأسلوب يعتمد على الوصف والتفسير بدلاً من القياس والتنبؤ.

لقد طُبقَ هذا النموذج في الستينات وأوائل السبعينات من القرن الماضي في الولايات المتحدة وانكلترا بدراسة بعض الموضوعات التربوية المتعلقة بالطلبة وطرائق التدريس مثل: دراسة الحالة الاجتماعية للطلبة وثقافاتهم، وتقييمات الطلبة لموضوعات معينة، وأساليب التدريس المستخدمة ومشروعات دولية تتعلق بالكلفة والمنفعة في مجال تعليم الآباء في بعض الدول النامية (Stuffle- beam & Shinkfield, 1985).

لقد ركز هذا النموذج على مفهومين أساسيين في العملية التربوية هما:

النظام التعليمي Instructional System وبيئة التعلم Learning Inviro- ment، فالنظام التعليمي يشمل الفضاء التعليمي الواسع المكون من شبكة من العلاقات المتداخلة المتفاعلة التي تكون الكل المتكامل. أما بيئة التعلم فهي البيئة الاجتماعية والنفسية والثقافية والمؤسسية التي يعمل فيها المعلمون والطلبة

والإدارة، والتي تتفاعل بطريقة معقدة لتؤدي نمطاً فريداً من الظروف والضغوط والعادات والآراء وأنماط العمل التي تؤثر في عمليات التعليم والتعلم. إن تركيز هذا النموذج على هذين العنصرين الهامين في العمل التربوي وما ينشأ عنهما من عوامل فرعية عديدة متداخلة، يجعل تناول تقويم البرنامج بهذا الأسلوب النوعي ضرورة لازمة حيث تكون نشاطات هذا النموذج موجهة نحو الفاعلية والكلفة والنواتج المرجوة والقيم السائدة في النظام ككل أو في بعض أجزائه، وبذلك يسعى المقوِّم إلى وضع قائمة وصفية للبرنامج المقوَّم، أي أن التقويم يستبدل النظره المجرده للنظام بوصف تفصيلي لعمليات الواقع الفعلي بكل أبعاده.

أما المراحل الرئيسية التي تميز التقويم التنويري فهي كما يأتي:

1- **الملاحظة**: وتتم من خلال الزيارات الميدانية المتكررة التي يقوم بها المقوِّم للبرنامج أو المشروع التي تجعله على اتصال مباشر ومستمر بالقضايا والمشكلات التي تبرز من خلال العمل، أو المناقشات التي يُثيرها مع العاملين في البرنامج ويجمع المعلومات الناشئة عنها.

2- **الاستقصاء**: وفيها يحاول المقوِّم أن يختار موضوعات مهمة من البرنامج يجري عليها مزيداً من البحث والتمحيص.

3- **التفسير**: ويعني إعطاء تفسير معمق وشامل للظواهر التي لاحظها المقوِّم أو التي ركز عليها الضوء من خلال البحث من أجل التوصل إلى استنتاجات ومبادئ عامة تحكم تنظيم البرنامج المقوَّم وتفيد في تحديد أنماط علاقات السبب بالنتيجة في عمليات البرنامج ووضع النتائج في إطار تفسيري وصفي.

ومن الجدير ذكره أن المراحل السابقة تكون متداخلة ومترابطة ويتم فيها الانتقال من مرحلة إلى أخرى كلما اتضحت المشكلات أثناء عمليات التنفيذ، ومن ثم تتم صياغة المعلومات المستخلصة من معلومات كثيرة توفرت لدى الباحث، في تقرير نهائي واضح تسهل قراءته وفهمه يشمل الموضوعات الرئيسية والتساؤلات الهامة المثارة حول البرنامج (علام، 2003).

ب. نموذج التقويم الطبيعي **Naturalistic Evaluation Model**

يُشير جوبا ولنكولن (Guba & Lincolen, 1981) إلى أن مفهوم البحث أو التقويم الطبيعي قدمت له تعريفات عدة أهمها ثلاثة، فهاوس (House, 1976) حدد تعريفاً للتقويم الطبيعي بأنه التقويم الذي يحاول التوصل إلى تعميمات طبيعية على مجموعة من الناس ولاسيما الناس غير التقنيين كالمعلمين والعامة من المجتمع باستخدام اللهجة العامية المبنية على تحليلات غير رسمية واستخدام حجج منطقية تحاول تأسيس بناء حقيقي للواقع المدروس. أما وولف وتايمتز (Wolf & Tymitz, 1977) فيعرفانه بأنه الاستقصاء الذي يهدف لفهم الوقائع الاجتماعية والتصورات الإنسانية التي ترصد دون تلوثها بالمتغيرات الجانبية للقياس الشكلي أو الأسئلة ذات التصور المسبق، بل أنه أسلوب يعتمد على أساليب قصصية يتم الحصول عليها من الأفراد مباشرة عن حوادث واقعية وقضايا طبيعية. أما وليامز وروش (Willems & Raush, 1969) فيعرفان التقويم الطبيعي بأنه الأسلوب الاستقصائي المعتمد على قدرة الباحث على توثيق بُعدين هما: ظروف الحوادث السابقة، ونتائجها. ويهدف أسلوب التقويم الطبيعي Naturalistic Evaluation إلى تنقية وتحسين شكل البرنامج المقوَّم وزيادة فاعليته وإدارته من خلال فحص عمليات تنفيذه أو تعديله بما يناسب منطقة معينة أو الأفراد المستفيدين من البرنامج من خلال تحسين الأساليب الفنية والخدمات التي يقدمها للأفراد المقيمين في منطقة معينة. ويتعامل هذا النموذج مع المواقف كما تحدث في الواقع وهي في بيئتها الطبيعية وبذلك يكون أسلوباً إستكشافياً لا يضع قيوداً معينة على النواتج ويقلل من تأثير المقوِّم في مكان التقويم ويركز على العمليات والتأثيرات الفعلية للبرنامج في فترة زمنية محددة. ويرى البعض أن هذا النموذج يستخدم الأسلوب الكلي (الجشتالتي) في فحص الواقع كما يجري في صورته الطبيعية، ومن خلال تفاعله مع الأطراف المستفيدة والمسؤولة عن البرنامج بأسلوب تسوده الثقة والود وبذلك يرسم صورة طبيعية واقعية لأنشطة البرنامج وفهم كامل لأنشطتهِ الإنسانية والمادية في إطارٍ طبيعي (علام، 2003).

ويستخدم هذا النوع من التقويم خطوات عدة تبدأ بـ:

أ. جمع البيانات من الموقع الطبيعي المستهدف تقويمه.

ب. تحليل وتقييم تلك البيانات.

جـ. إثارة أو وضع أسئلة نوعية يسترشد بها المقوِّم بشكل مستمر.

د. العودة إلى موقع البرنامج كلما دعت الحاجة للمزيد من البيانات.

وخلال هذه الدورة من خطوات التقويم يتم وضع فروض وتخمينات ذكية يتم اختبارها في ضوء البيانات التي يتم جمعها والتحقق من صحتها (الفروض) أو إعادة صياغتها إذا تطلب الأمر أو رفضها. ويعتمد المقوّم وفق هذا الأسلوب على الملاحظات المتتالية ويدوِّن وقائعها، ويجري مقابلات يضع لها خططاً مسبقة ويجري حوارات مفتوحة مع الأفراد العاملين أو المستفيدين من البرنامج، وقد يستخدم لتسجيل هذه البيانات أجهزة تسجيل صوتية أو مرئية، ويطلع على الوثائق الرسمية الخاصة بالبرنامج وآراء وانطباعات ما يقوله المشاركون في البرنامج ويقوم بتحليلها، ويركز على التناقضات (إن وجدت) في تلك الملاحظات والأقوال والسجلات الرسمية ليستنتج منها بعض الأفكار المفيدة التي تساعده للتوصل على استنتاجات وحقائق موضوعية. ليس هذا فحسب بل أنه ومن خلال ما تقدم قد تتولد لديه تساؤلات جديدة توجه البحث وجهة أخرى قد لا تكون محددة سابقاً أي عند بداية التقويم، وعند ذاك يعود ثانية إلى الموقع لتوفير إجابات لتلك التساؤلات. وأخيراً لابد من الإشارة إلى أن التقرير النهائي للتقويم وفق هذا الأسلوب يعتمد على الوصف والتفسير.

تعقيب وتعليق على نماذج التقويم:

قدمنا للقارئ في الصفحات السابقة عرضا للنماذج المشهورة في مجال تقويم البرامج للاستفادة منها بعد التعرف على منطلقاتها النظرية والأساليب الإجرائية لاستخدامها، إلا أن من الضروري جداً للمقوم أن يحاول وقبل اختيار النموذج المحدد أن يوازن ويحدد العلاقة العضوية بين منطلقات النموذج الذي يختاره وطبيعة وواقع البرنامج الذي يريد تقويمه ويحاول أن

يضع خطة متصورة بسيطة لكيفية الاستفادة من النموذج لتقويم البرنامج المطلوب، فإن توصل إلى وجود نقاط التقاء عديدة واضحة بينهما، وحتى إن وُجدَت نقاط أخرى قليلة متعارضة فإن بإمكانه تكييفها لخدمة عملية التقويم أو استثنائها من العملية متذكراً أنه لا توجد وصفة دوائية واحدة لعلاج جميع الأمراض، وبلا شك فإن خبرة وفهم المقوِّم لطبيعة نماذج التقويم من جهة والبرنامج المقوَّم من جهة أخرى سيسرع في عملية الاختيار الأفضل. إن النماذج التي ذكرت تتباينً في أساليب تطبيقها وفي تحديد المعايير التي يجري في ضوئها إصدار الأحكام عن البرنامج المقوَّم.

فهناك المجموعة التي تعتبر الأهداف معايير أو محكات ومن خلال تحقيقها كلياً أو جزئياً يمكن الحكم على نجاح أو إخفاق البرنامج، بينما ينتقد البعض من واضعي النماذج الاخرى هذا الأسلوب ويعتبروه غير كافٍ لاسيما وأن الأهداف نفسها لم توضع إلا لتخدم أغراض المسؤولين عن البرامج وليس المستفيدين منها، مع العلم أن هذه النماذج تستخدم بشكل كبير في العديد من دول العالم لاسيما الأقطار العربية ولاشك أن صعوبة صياغة معايير أو محكات موضوعية دقيقة هي التي أعطت الفرصة لانتشار هذا النوع من النماذج. أما نموذج سكريفن Secrevin والآخرون من مؤيديه فعلى الرغم من أهميته المتجسمة في التركيز على الأثر الذي يتركه البرنامج في تغيير البيئة التي وجد من أجلها، والاعتماد على آراء المستفيدين منه، إلا أن تطرفه في رفض الاعتماد على الأهداف بل وحتى النظر إليها يتضمن نوعاً من التطرف، لاسيما إذا ما علمنا أن ما يريده سكريفن من البرنامج قد يكون من أولويات أو أهداف البرنامج الذي يرفض هو النظر لها، وبذلك يمكن أن نفقد الكثير من الفوائد التي تجنيها عملية التقويم من تلك النماذج.

أما بالنسبة للنماذج التي تهدف لمساعدة متخذي القرار فتبدو اساليب موضوعية لاسيما وأن من أولويات عمليات التقويم هو التوصل لاتخاذ قرارات صائبة عن البرنامج المقوّم وبذلك تقدم هذه النماذج مصدراً مساعداً

هاماً لتسريع هذه العملية، ليس هذا فحسب بل أن من بين تلك النماذج ذلك الذي يدعو إلى اعتماد معايير موضوعية عالمية موحدة للحكم على البرامج والمؤسسات التعليمية من خلال مؤسسات الاعتمادية الدولية وما تتضمنه هذه العملية من مساعدات علمية يقدمها خبراؤها لتلك المؤسسات من نصائح وتوصيات وإرشادات تساعدها في ردم مواطن الضعف بأساليب علمية مجربة ومن ثم إعادة فحص ما تم عمله من إصلاحات قبل الموافقة على إعطائها الاعتماد بشكله الرسمي، ولاسيما أن هذا الأسلوب يقدم خدمات إيجابية لتطوير البرامج من خلال تقويمها المتنوع والمستمر رغم كلفته المادية العالية التي قد لا تستطيع بعض البلدان الإيفاء بها، وبذلك تحرم من فوائده.

أما الأساليب النوعية الطبيعية فتتميز بشكل عام بميزتين هما: قلة الاعتماد على الأساليب الكمية سواء في جمع أو تحليل البيانات على أساس أن هذه الأساليب قد لا تعكس حقيقة المشاعر والاستجابات الصحيحة وهي قابلة للتحريف والتزييف، أما الميزة الثانية في تلك النماذج فَهي اعتمادها على الملاحظات الطبيعية ورصد الظاهرة قيد الدرس في وضعها الطبيعي الذي يعكس الواقع كما هو دون إضافة أو تحريف. ولاشك أن هذا الأسلوب رغم التعقيدات المصاحبة له من حيث الاعتماد على نقل وتصوير الواقع كما هو من ناحية، أو اعتماد الباحث على خبراته وانطباعاته الشخصية في تحليله للبيانات أو تفسيرها وما يصاحب ذلك من انعكاس العوامل الذاتية على تلك النشاطات التي يقوم بها المقوِّم، لكن المقوِّم الواثق بنفسه، الذي يملك من الخبرات الكثيرة في المجال المقوَّم، ذو المعرفة الجيدة لأساليبَ البحث والتقويم قد يقلل الكثير من تلك التخوفات، ويزيد هذا التقويم موضوعية أكثر إذا ما شارك في عملية التقويم أكثر من مقوِّم.

وعلى الرغم من عدم وجود إتفاق كامل بين الخبراء حول أفضل نموذج يمكن الأخذ به، فإن من نافلة القول أن التقويم يجب أن يتضمن أسس عامة أهمها قيام المَقوِّم بوضع خطة تفصيلية عما سيقوم به خلال عملية التقويم تتضمن الأهداف المتوخاة منه والتصميم أو النموذج الذي سيسترشد به خلال

التقويم وتصوراً كاملاً لأدواته وأساليب تحليل البيانات وتقديم الاستنتاجات والتوصيات بشكل إجرائي قابل للتطبيق ويتماشى مع طبيعة البرنامج والإمكانيات المتاحة للتطوير أو التغيير، وأن يفهم المقوِّم أو يتعرف على الأطر الفكرية للمسؤولين عن البرنامج وأن يستخدم لغة واضحة ومبسطة للتعرف على مشاعر المستفيدين ورغباتهم وآرائهم في تطوير عمل البرنامج إلى شكل أفضل أو إيقافه عن العمل. وأخيراً لابد من طرح مجموعة من التساؤلات المطلوب من المقوِّم الإجابة عليها ومنها:

1- هل التقويم الذي يقوم به الباحث يهدف للوصف، أم إصدار الأحكام، أم هما معاً، أم مساعدة المسؤول في اتخاذ القرار المناسب؟

2- هل الدراسة التقويمية تهتم بالسوابق Antecedents والتفاعلات Transac-tions أم أنها تركز فقط على النواتج Output؟

3- على المقوِّم أن يحدد أي من الأسلوبين الكمي أم النوعي يتم استخدامه في التقويم؟ وهل يمكن جمع الأسلوبين معاً. لأن ذلك سينعكس على تحديد أهداف التقويم والأدوات التي سيستخدمها لجمع بياناته.

4- هل يهدف التقويم إلى مطابقة واقع البرنامج مع أهدافه لتحديد جوانب الخلل والقوة وتقديم الاقتراحات؟ أم يحاول وضع معايير موضوعية معينة، ويطابق بينها وبين واقع البرنامج من أجل تحديد جوانب الضعف أو القوة؟

5- هل يقتصر التقويم على واقع البرنامج بذاته، أم بمقارنته ببرامج أخرى مشابهة، لتحديد جوانب الإخفاق أو النجاح؟

6- هل سيقدم المقوِّم تقريره عن البرنامج إلى المسؤول عنه مباشرةً، أم إلى الجهات الأعلى التي كلفته بالقيام بهذه المهمة؟

مراجع الفصل الثالث

1- الدوسري، إبراهيم بن مبارك (2000). الإطار المرجعي للتقويم التربوي. الرياض: مكتب التربية العربي لدول الخليج.

2- الشبلي، إبراهيم مهدي (1984). تقويم المناهج باستخدام البرامج. بغداد: مطبعة المعارف.

3- علام، صلاح الدين محمود (2003). التقويم التربوي المؤسسي: أسسه ومنهجياته وتطبيقاته في تقويم المدرس. القاهرة: دار الفكر العربي.

4- الصارمي، عبد الله بن محمد (2003). تقويم المنهج المدرسي. بحث غير منشور.

5- Bodgan, R. C., & Biklen, S. K. (1982). **Qualitative Resrearch for Education**. Boston: Allyn & Bacon.

6. Campbell, D., & Stanley, (1966). **Experimental and Quasi- Experimental Designs for Resrearch**. Chicago: Rand Mc Nally.

7. Cook, T.D, & Reichardt, C. S, (Ed), (1979). **Qualitive and Quantitative Methods in Evaluation Research**. California: SAGE Publications, Inc.

8. Dickey, F. G. & Miller, J. W. (1972). **A Current Perspective on Evaluation**. Washington Dc. American Association for Higher Education.

9- Fistzpatrik, J., & Worthen, B. (2004), **Program Evaluation: Alternatie Approaches and Practical Guidines**. New York: Allyn and Bacon.

10. Guba, E. G., & Lincoln, Y. S. (1981). **Effective Evaluation.** California: Jossey- Bass Publishers.

11. Hamilton, D., & Others (1977). **Beyond the Numbers Game**. California: McCutchan.

12. House, E. R. (1976). **Justice in Evaluation**. In G. V. Glass (Ed.). Evaluation Studies Review Annual. Vol.1. Beverly Hills, California, SAGE, Publications.

13. Parlett, M. (1977). **Introduction to Illuminate Evaluation**. California: Pacific Sounding Press.
14. Sanders. J.R., & Sachse, T. P. (1977). **Applied performance testing in the classroom**. Journal of Research and Development in Education, 10, 92-104.
15. Scriven, M. (1973). **Goal Free Evaluation**. In E. R. House (Ed.), School Evaluation: The Politics and Process. Berkeley, Calif: McCutchan.
16. Scriven, M. (1984). **Evaluation Idiologies**. In R. F. Conner, D. G. Altman, & C. Jackson (Eds), Evaluation Studies Review Manaul (Vol.9) Beverly Hills, CA: SAGe.
17. Stake, R. E. (1967). **The Countenance of educational evaluation.** Teacher College Records, 68, 523-540.
18. Stufflebeam, D. L. (1973). **Experts from "Evaluation as enilghtment for desision making"** in B. R. Worthern & J. R. Sanders,(Ed). Educational Evaluation: Theory and Practice. Belmont, CA: Wadsworth.
19. Stufflebeam, D. L., & Shinkfield, A. J. (1985). **Systematic Evaluation**. Boston: Kluwer nijhoff Publishing.
20. Willems, E. P. & Raush, H. L. (1969). **Naturalistic Viewpoints in Psychological Research**. New York: Holt, Rinehart and winston.
21. Wolf, R. L., & Tymitz, B. (1977). **Toward More Natural Inquiry in Education**. CEDR Quarterly. 10, pp. 7-9.
22. Worthern, B. R., & Sanders, J. R. (1984). **Content Evaluation and Educational Evaluation: Anecessary Message**. Kalamazoo: Western Michigan University, Evaluation Center.
23. Worthen, B. R., & Sanders, J.R (1973). **Educational Evaluation: Theory and Practice**. California: Wadsworth Publishing Company, Inc.

الفصل الرابع

تخطيط التقويم ومعاييره

تخطيط التقويم

كتابة التقرير النهائي للتقويم

توصيات لزيادة استخدام نتائج التقويم

العلاقة بين المقوم ومتخذ القرار

معايير ومحكات التقويم التربوي

4



الفصل الرابع

تخطيط التقويم ومعاييره

تخطيط التقويم

ليس هناك شك في أن أي عمل ينوي الإنسان تنفيذه لابد وأن تسبقه عملية تخطيط تحدد طبيعة الشيء المراد عملهُ، والمستلزمات والأساليب الواجب استخدامها في عملية التنفيذ، وأن التخطيط للشيء قبل تنفيذه يختصر الزمن ويقلل الجهد والوقوع بالخطأ، ويوصل إلى تحقيق الهدف المنشود.

والتقويم شأنه شأن العمليات الإدارية الأخرى تحتاج إلى المقوِّم أن يحدد تفاصيل الخطة التي سيقوم باتباعها لتنفيذ عملية التقويم ودراستها ومراجعتها عدة مَرات، ومن ثم عرضها على بعض الخبراء في مجال التقويم والاستفادة من ملاحظاتهم في مراجعتها وتعديلها.

ليس هذا فحسب بل من الضروري في بعض الأحيان مناقشة خطة التقويم مع المسؤول عن البرنامج المقوَّم واحترام آرائه وملاحظاته والاستفادة منها، إضافةً إلى الاتفاق معه حول موعد البدء بتنفيذ عملية التقويم والمستلزمات المادية والبشرية التي يحتاجها المقوّم لتسهيل مهمته كإجراء المقابلات اللازمة للحصول على معلومات ضرورية لعملية التقويم.

لقد أورد (Worthen & Sanders, 1973) مقترحات وتوضيحات نظرية وعملية لكيفية وضع خطة تقويم البرنامج وكتابة تقريره النهائي على المقوّم الأخذ بها ومنها:

أولاً: توضيح لماذا ستتم عملية التقويم: أي بيان الاسباب الموجبة لاجراء عملية التقويم.

ثانياً: تحديد الأفراد الذين يُفترض أن يزودهم بالمعلومات التقويمية وتحديد حاجاتهم وخلفياتهم حتى يتسنى له صياغة تقرير التقويم ونتائجه بشكلٍ يناسب مستوياتهم وخلفياتهم، فمثلاً إذا ما أريد منه أن يقدم ذلك

التقرير إلى أولياء أمور الطلاب فهو بالتأكيد سيكتبه بأسلوب يختلف عنه لو طلب إليه أن يقدم التقرير للمسؤولين الحكوميين عن البرنامج.

ثالثاً: على المقوِّم أن يضع خطة شاملة للتقويم تتضمن التصميم، مواصفات البيانات التي سيقوم بجمعها خلال عملية التقويم، أساليب جمع وتحليل البيانات ووصفاً لكيفية استخدام المعلومات والبيانات التي سيم الاعتماد عليها في عملية التقويم.

هذه الخطوات المقترحة هي أمثلة من بين أمور أخرى عديدة يجب أخذها بعين الاعتبار في تخطيط عملية التقويم يوصي بها ماكنمارا (McNamara, 1998) منها:

1- تحديد الغرض الذي نريد تحقيقه من عملية تقويم البرنامج، أي ماذا نريد عمله بنتائج التقويم وما القرارات الممكن اتخاذها بناءً على تلك النتائج.

2- تحديد ماهية المعلومات التي نريد جمعها لاتخاذ قرار ما؟ هل تلك المتعلقة بـ (نواتج البرنامج)، أم المتعلقة بـ (المدخلات والعمليات والنواتج) معاً، أو مدى استفادة الأشخاص المنتفعين من البرنامج؟ أم أننا نريد من التقويم تركيز الانتباه على نقاط الضعف فيه، أو التركيز على الجوانب التي فشل في تحقيقها للمستفيدين منه؟

3- من هم الذين سنقدم لهم نتائج التقويم؟ هل هم المستفيدين أم الممولين للبرنامج، أو الجهاز الإداري المسؤول عنه، أم العاملين في البرنامج.

4- ما هي المصادر التي يجب أن تجمع منها المعلومات المطلوبة؟ العاملين في البرنامج، المستفيدين، أم الأثنين معاً؟

5- ما الطرق المثمرة والأساليب التي تجمع المعلومات بواسطتها؟ الاستبيان، المقابلات، فحص الوثائق، ملاحظة العاملين وغيرها.

6- ما الوقت المناسب لجمع المعلومات، أثناء العمل أم بعده؟

7- هل مصادر جمع المعلومات متوفرة خلال التقويم؟

ويضع كل من (Worthen & Sanders, 1973) استمارة لمخطط تقويمي مقترح يتضمن:

1- المبررات: وتوضح من خلالها الأسباب التي تدعو للقيام بعملية التقويم.

2- تحديد أهداف الدراسة التقويمية. ويتضمن ذلك الآتي:

أ. ما النتائج المتوقعة لعملية التقويم؟

ب. من هم الأفراد الذين ستخدمهم الدراسة التقويمية؟

3- وصف البرنامج الذي سيتم تقويمه، من حيث:

أ. فلسفة البرنامج.

ب. محتوى البرنامج.

جـ. أهداف البرنامج الواضحة والضمنية.

د. وسائل البرنامج (الاستراتيجيات، الأعلام. .الخ).

هـ. الطلبة.

و. المجتمع الذي يقع ضمنه البرنامج والبيئة التعليمية للبرنامج.

4- التصميم المستخدم في التقويم:

أ. محددات تصميم التقويم.

ب. الخطة التنظيمية العامة (أو النموذج) الذي سيتم بموجبه تقويم البرنامج.

جـ. الأسئلة التقويمية.

د. المعلومات المطلوبة للإجابة عن الأسئلة.

هـ. مصادر المعلومات، أساليب جمع المعلومات.

و. جدول ينظم عملية جمع المعلومات.

ل. أساليب جمع وتحليل المعلومات.

م. المعايير التي يتم الاعتماد عليها في الحكم على النوعية.

ن. أساليب كتابة تقرير التقويم.

ي. الميزانية المقترحة لإنجاز عملية التقويم.

5- وصف التقرير النهائي:

أ. مخطط التقرير الذي سيقدم من قبل المقوِّم.

ب. الفائدة المتوخاة لنتائج الدراسة التقويمية.

جـ. إدراك تحيزات المقوِّم التي يمكن أن تؤثر على محتوى التقرير النهائي للتقويم، إذ هناك احتمال أن المقوِّم قد يحمل نظرة سالبة أو موجبة عن البرنامج قبل أن يكلف بعملية التقويم ومن المحتمل أن يحملها معه في مراحل التقويم أو حتى عند كتابة التقرير النهائي لاسيما أن المقوّم هو كائن بشري وقد لا يكون دائماً موضوعياً بشكل كامل ولذلك عليه أن يتذكر هذه الحقيقة دائماً ويحاول أن يبتعد عن تحيزاته الذاتية ويحقق أعلى قدر مستطاع من الموضوعية.

أما (Psavac, and Carey, 1985) فقد اقترحا ست خطوات إعتبراها ضرورية لتخطيط عملية تقويم البرامج نوجزها كما يأتي:

خطوات تخطيط التقويم

الخطوة الأولى: تحديد المعنيين بالتقويم

ان أول ما يجب أن يفعله المقوم الناجح هو أن يحدد المعنيين بالتقويم أو أولئك الذين لديهم اهتمام جاد بالبرنامج محل التقويم والذين قد يتأثروا بعملية التقويم، وهؤلاء قد يكونوا من العاملين بالبرنامج والذين يحصلون على دخلهم أو جزء منه من العمل في البرنامج ، والذين يمكن أن يتأثر مستقبلهم الوظيفي بنتائج التقويم، كما يمكن أن يكونوا المستفيدون من الخدمات التي يقدمها البرنامج.

ويجب مراعاة العاملين في البرنامج أولاً لارتباطهم به إذا ما قورنوا بمولي البرنامج أو المستفيدين منه، فمدير البرنامج يعتبر الشخصية المحورية الذي سيعتمد عليه المقوم طوال المشروع، ولذلك يجب معرفة كل ما يمكن عن خلفيته العلمية، وميوله ورغباته لأنها قد تسهم اسهاماً ايجابياً في عملية التقويم. كما يجب عدم إغفال القائمين على تقديم خدمات البرنامج وأهمية إشراكهم في مرحلة التخطيط ليتبنوا مشروع التقويم وتقديم كل الدعم للمقوم

أثناء مرحلة جمع البيانات، كما أن من الضروري أن يفهم المقوم طبيعة العلاقة التي تربط مدير المشروع بالعاملين فيه.

يأتي بعد ذلك مباشرة مراعاة مولي البرنامج او المسؤولين عن توفير الدعم المالي للبرنامج، ففي بعض الأحيان العاملون في البرنامج هم أنفسم ملاك البرنامج وفي أحيان أخرى ليس كذلك. وعادة ما يكون الرعاة مؤسسات الدعم، أو أجهزة حكومية أو مدراء المؤسسة التي يطبق فيها البرنامج، وعادة ما يكون هناك شخص واحد يرجع إليه المقوم وفي أحيان أخرى أكثر من شخص، إن التواصل مع رعاة البرنامج له أهميته وخاصة في مرحلة التخطيط لتقويم البرنامج وفي نهاية عملية التقويم، ففي مرحلة التخطيط يجب التأكد من دعمهم الكامل لعملية التقويم المقترحة، وفي نهاية عملية التقويم لمعرفة ملائمة التقرير النهائي لهم في عملية اتخاذ القرار.

وأخيراً لا بد من تحديد المستفيدين من خدمات البرنامج، فحجم ونوع التواصل مع المستفيدين من البرنامج يعتمد على طبيعة البرنامج وعملية التقويم معاً، فعلى سبيل المثال إذا ما أردنا تقويم أداء مدرسة معينة فبالإضافة الى اهتمامنا بآراء المدير العام للمنطقة التعليمية التي تقع فيها المدرسة كمسؤول أعلى ومديرها ومعلميها فإن الاهتمام بالمستفيدين منها آباء وطلاب يعتبر من اوليات عمل المقوم من خلال عقد لقاءات معهم والاستماع إلى آرائهم وملاحظاتهم التي تعتبر ضرورية قبل البدء بعملية التقويم كأهميتها إثناءه وفي نهايته.

الخطوة الثانية: عقد اجتماع تحضيري

قبل الاتفاق على القيام بعملية التقويم، وقبل الشروع في كتابة مخطط تفصيلي لها، لعل من المناسب للمقوم الاجتماع بكافة المعنيين بعملية التقويم للوقوف على بعض المعلومات الأولية حول الأسئلة الخمسة التالية:

(1) من الذي يرغب في عملية التقويم؟

(2) ما نوعية التقويم المطلوب؟

(3) ما الدوافع وراء التقويم؟

(4) ما الإطار الزمني المتاح؟
(5) ما هي المصادر والتسهيلات المتاحة؟

من الذي يرغب في عملية التقويم؟ الوضع المثالي يحدث عندما يكون كل من رعاة البرنامج والعاملين فيه يرغبون في قيام عملية التقويم، في هذه الحالة سيتعامل المقوم مع أناس ثقات يؤمنون بخبراتهم العملية ومنفتحين لمقترحات التطوير، ومرحبين بتوثيق وجهات نظرهم حول البرنامج. أما حين يشرع راعي البرنامج في عملية التقويم سواء بدون العاملين أو غصبا عنهم، فان ذلك يتطلب من المقوم القيام بإقناع العاملين في البرنامج بأهداف ومنهجية التقويم قبل الشروع في عملية جمع البيانات، وعند عدم نجاحه في ذلك، فقد يواجه احتمال معارضتهم أو عدم تعاونهم معه، بمعنى آخر عندما يرى العاملون في البرنامج بان المقوم حليفاً وليس عدوا لهم يكونوا أكثر تعاوناً في عملية جمع البيانات وتقديم التفسيرات المعمقه لتلك البيانات.

عندما يشرع العاملون في البرنامج في عملية التقويم بدون علم راعي البرنامج أو ضد رغبته فان ذلك يتطلب من المقوم إقناع رعاة البرنامج بأهمية التقويم وإلا يصعب عليه تحقيق تقدم يذكر. فالرعاة الذين لا يتبنون عملية التقويم أثناء عملية التخطيط لها يميلون الى جعل تقرير التقويم حبيس أدراجهم.

ما نوعية التقويم المطلوب؟ في الاجتماعات الأولى مع راعي البرنامج أو القائمين عليه عادة لا يكون لمصطلح "تقويم البرنامج" نفس المعنى عند الجميع. فهو يعني في اغلب الأحيان للقائمين على البرنامج أنه تقويم تكويني سيساعدهم على تعديل وتحسين البرنامج. في حين يرغب رعاة البرنامج في التقويم الختامي لأنهم عادة ما تمارس عليهم الضغوط لتوجيه المصادر المالية إلى برامج أخرى ويجب عليهم تقرير استمراريتهم في دعم البرنامج من عدمها. وأخيراً فان بعض العاملين في البرنامج ستكون لديهم معلومات ضحلة عن تقويم البرامج ويتوقعون نوعا من التقييم الخاص لهم.

وهنا يجب على المقوم توضيح مفهوم التقويم لأولئك الذين لا يفهمونه

ومساعدتهم لاختيار نوع التقويم الذي يتماشى مع رغباتهم وحاجياتهم وإمكانياتهم. ونادراً ما يكون الخيار بين جملة من العناصر من أنواع متعددة من التقويم تعتمد على حجم البرنامج وتشعب أهدافه والإمكانيات المادية والبشرية المتوفرة له.

لماذا الرغبة في التقويم؟ هذا السؤال يرتبط بسابقه. ان طلب القيام بالتقويم لم يكن أبداً ناتجاً عن الرغبة في روح البحث العلمي وحدها، ولكنها غالباً ما تكون استجابة لدوافع سياسية. فرعاة البرنامج يهمهم الإبقاء على البرنامج الذي يظهرهم بالمظهر الجيد أمام المجتمع ورؤسائهم. التقويم الفعال إذاً يعطي أهمية قصوى لمعرفة الأسباب وراء الرغبة في عملية التقويم. هل هناك جماعات معينة تعترض على القيام بعملية التقويم؟ وما هي دوافعهم؟ وهل هناك نية صادقة من قبل العاملين في البرنامج ورعايته للاستفادة من نتائج التقويم لتحسين عملية اتخاذ القرار؟ ان ما يهم العاملين في البرنامج هو البحث عن إجابة لسؤال ملح عن مستقبل البرنامج، هل سيستمر؟ هل سيتوسع؟ هل سيتم تعديله؟

ان أحد وظائف التقويم مساعدة كل من رعاة البرنامج والعاملين فيه للوصول الى قرار أو قرارات بالإجماع حول البرنامج. وهذه القرارات قد لا تكون متعلقة بمدى فائدة البرنامج فبعض البرامج فوائدها واضحة للعيان وإنما تتعلق بجوانب أخرى أكثر تعقيداً.

وفي حالات أخرى يقام التقويم لتحديد أثر البرنامج على المؤسسة الراعية وتأثرها على النواحي الأخرى للمؤسسة. ففي هذه الحالة تكون القرارات غير موجهة نحو القائمين في البرنامج.

كما إن بعض عمليات التقويم تجرى لأغراض سياسية أو من أجل الرأي العام، فعلى سبيل المثال يمكن أن يكون مُديرو البرنامج على يقين من أن لديهم برنامجاً فعالاً ولكنهم يطمحون لمزيد من الدعم لبرنامجهم من خلال تسويق نتائج التقويم لأصحاب القرار.

ويمكن أن يكون التقويم كأحد متطلبات الدعم المالي. فمعظم المؤسسات المانحة للدعم حكومية كانت أم خاصة تطلب نوعاً من التقويم. وفي هذه الحالة فإن قرار التقويم ذو منشأ خارجي عن البرنامج نفسه. أن فرض عملية التقويم لا تخلو من حسنات. فالأسئلة التقويمية عندما تفرض على العاملين في البرنامج تدفعهم للنقد الذاتي أكثر من الحالات الأخرى. ويجب أن نتذكر بأن الحاجة قائمة للحصول على دعم للبرنامج الجيد من اجل تبرير استمرار الداعمين لتمويل البرنامج. ويبقى القول بان هناك ثمة أسباب للتقويم غير مقبولة أو محبذة، فعلى سبيل المثال إداريو البرنامج يمكن أن يستغلوا تقويم البرنامج كتسويق لاتخاذ قرارات مهمة حول البرنامج. كما إن الإداريين الذين يشعرون بالضغط على مدى جدوى البرنامج يمكن أن يتعللوا بأن البرنامج تحت تقويم، وان عملية التقويم تأخذ وقتاً طويلاً، وبتقديم دعم محدد أو ناقص للمقومين يضمنوا بأن التقويم سيأخذ وقتاً أطول للانتهاء منه، وإذا ما بدت الحاجة لكسب مزيد من الوقت بعد الانتهاء من عملية التقويم فلن يترددوا من تعيين لجنة لدراسة نتائج التقويم. كما أن التقويم ليس مناسباً عندما يعرف الإداريون القرارات التي يجب اتخاذها ويخوضوا لعبة التقويم لإضفاء الشرعية على قراراتهم لا غير .

ما التوقيت المناسب للتقويم؟ مجال عملية التقويم وعمقها تحددها العديد من العوامل، فإذا ما كان الوقت المتاح للمقومين ضاغطاً فان ذلك من شأنه أن يرغمهم لاتخاذ خيارات قد لا تكون مثالية، فمثلاً يمكن أن يلجأ القائمون بعملية التقويم الى استبانات أو أي أداة أخرى غير معروف مدى صدقها أو ثباتها أو قابليتها للتطبيق بدلا من أدوات مصممة خصيصاً لموضوع التقويم، أو كأن يكون عدد أفراد العينة قليلاً مما يقلل من احتمالية الكشف عن الأثر الحقيقي للبرنامج. ففي مثل هذه الأحوال غير المثالية لعملية التقويم يجب على المقومين اتخاذ القرار المناسب حول المضي قدماً في عملية التقويم، أو تأخيرها أو رفض القيام بها حتى لا يتوصلوا إلى نتائج مضللة. وحتى عندما

يكون هناك متسعاً من الوقت، يمكن أن يقوم القائمون على البرنامج بدافع عجلتهم للحصول على النتائج بالضغط على المقومين لتحديد وقت للانتهاء من عملية التقويم. فمن المستحسن عدم الالتزام بتحديد وقت للانتهاء من عملية التقويم ويمكن الالتزام به فقط في حالة أن كل الأمور سارت على ما يرام. فهناك العديد من الأمور التي يمكن أإن تكون خارج السيطرة كأن يمرض جامعو البيانات، أو احتمال صعوبة الحصول على المستجيبين، أو تعطل الحاسب الآلي قي وقت تكون الحاجة إليه قائمة.

عندما يكون تحديد موعد للانتهاء من التقرير النهائي للتقويم أمراً مطلوباً يجب أخذ متسع من الوقت. فإذا ما انتهى المقومون من عملية التقويم بعد الوقت المحدد لها فهم مهتمون في أحسن الأحوال بأنهم غير منظمين وفي أسوأ الأحوال غير أكفاء. في المقابل عندما ينتهون قبل الموعد المحدد فإنهم يبنون سمعة بأنهم يتسمون بالسرعة والكفاءة. وهنا يجب أن نتذكر بأن المقومين يعملون في اطر يلعب عامل الوقت فيها دوراً مهما تختلف عن السياق الذي يعمل فيه البحث الأكاديمي.

ما المصادر والتسهيلات المتوفرة؟ بجانب عامل الوقت هناك عامل آخر يمكن أن يحد من عملية التقويم الا وهو توفر المصادر وخاصة المالية منها، ومدى انشغال القائمين بالتقويم في مشاريع تقويمية أخرى.

الخطوة الثالثة: الحكم على قابلية البرنامج للتقويم

إن هدف هذه الخطوة هو تقرير ما إذا كانت عملية التقويم قابلة للتنفيذ، فبعد الاجتماع بالمعنيين بالبرنامج لجمع البيانات الأولية حول ماهية البرنامج وأهدافه ومصادره يجب تقرير قابلية البرنامج للتقويم من عدمه.

ولا يمكن الحكم على قابلية البرنامج للتقويم قبل الحصول على إطاره النظري، وتحديد المصادر اللازمة ولا عندما تكون التوقعات من البرنامج مستحيلة التحقيق، ولا عندما يستحال تجميع الأدلة والإثباتات على فاعلية

البرنامج. كما لا يمكن تقويم البرنامج إذا لم يكن تطبيقه كما هو مخطط له، فعلى سبيل المثال إذا ما خططت مدرسة لإدخال مقرر حاسب آلي للصف الثامن يتطلب مدرس متفرغ وعشرة حواسبيب، فان البرنامج لا يمكن تقويمه في حالة وجود ثلاثة حواسيب ومدرس غير متفرغ.

الخطوة الرابعة: مراجعة الأدبيات

كثيراً ما ينشغل المقومون في مشاريع متنوعة، فالمقوم الواحد يمكن أن يقوم- على سبيل المثال- نظاماً مدرسياً، وبرنامجاً مجتمعياً عن الإدمان وكذلك المناخ التنظيمي لدائرة ما، مثل هذا الباحث أو المقوم لا بد أن يطلع على مجموعة واسعة من الأدبيات. وبالرغم من صعوبة ذلك فان بعض الباحثين يحبذونه، انهم يحبذون التنوع على تركيز اهتمامهم في مجال معين.

ان المقومين الذين سبق لهم العمل في مجال معين يجدون صعوبة في تصميم خطة التقويم فإذا ما أنصب اهتمامهم في تقويم النظم التعليمية، وبرامج الصحة العقلية، أو المشاكل الشخصية في مؤسسة ما، فإنهم يبنون كماً معرفياً هائلاً عن التقويم في مجال تخصصهم، كما إنهم يستطيعون وبسهولة البناء على خبراتهم البحثية السابقة بحيث تسهم في بناء الإطار النظري في مجال التقويم المزمع.

ولكن عندما يشتغل المقومون في مجال جديد بالنسبة لهم، فمن الضروري أن يكونوا أكثر حرصاً في البحث في الأدبيات قبل تصميم أدوات بحث جديدة. إذ يجب عليهم الاستفادة من نجاحات وإخفاقات الآخرين والتوصل إلى صورة واضحة بالمصاعب المنهجية والسياسية والعملية التي يجب التغلب عليها.

ويمكن أن يكون البحث الألكتروني انطلاقة جيدة للبحث في الأدبيات المتعلقة بالموضوع. فهناك أربع قواعد بيانات قيّمة يمكن الانطلاق منها وهي:

- بيانات في تقويم البرامج .Data on Program Evaluation
- مد لاين .MED- LINE

- بحث واسترجاع ملخصات الدراسات النفسية Psychological Abstracts Search and Retrieval.
- مركز مصادر المعلومات التربوية Educational Resource Information Center (ERIC).

كما أن البحث في الدوريات التي يوصي بها العاملون في البرنامج يعد مفيداً فيعد الحصول على عدد قليل من الدراسات المفيدة وبالرجوع إلى قائمة المراجع فيها يمكن التوصل إلى مراجع إضافية.

من خلال قراءة الدراسات يجب على الباحثين وضع الاسئلة الآتية في الحسبان: هل تم تقويم برامج مماثلة سابقاً؟ ما المنهجية المستخدمة؟ هل تم تصميم مقاييس جديدة؟ ما مدى ثبات وصدق تلك المقاييس؟ ما نوعية التحليلات الإحصائية المستخدمة؟ وما مدى مناسبتها؟ هل هناك انسجام في نتائج الدراسات السابقة؟ أم أن هناك تضارب وهل سببه طريقة اختيار العينة؟ أم المنهجية أم تفسير النتائج؟ وما الجوانب التي لم يتم التطرق لها أو بحثها؟ ومن خلال ذلك يمكن للمقوم أن يحدد جوانب استفادته من تلك الأدبيات في مشروع التقويم الذي يخطط له.

الخطوة الخامسة: تحديد منهجية البحث التقويمي

بعد مراجعة الأدبيات يكون المقوم جاهزا لاتخاذ قرارات إجرائية بخصوص استراتيجية ومنهجية البحث التي ستستخدم في التقويم من حيث تحديد مجتمع الدراسة وطرق اختيار العينة، وأدوات التقويم، وجمع البيانات والتحليلات الإحصائية.

أ- المنهجية: منهجية البحث عادة ما تحددها نوعية التقويم المطلوب. فهل مجال التقويم المطلوب هو تقويم الاحتياجات Needs؟ أم العمليات Pro-cesses؟ أم النواتج Outcomes - أم فاعلية التكلفة Effectiveness Cost - علماً بأن التقويم يمكن أن يشمل أكثر من جانب من هذه الجوانب.

وتتأثر المنهجية أيضاً بعامل آخر يتعلق بما إذا كان البرنامج جاري تطبيقه أم انه ما يزال في مرحلة التخطيط فإذا ما كان البرنامج سيشرع فيه، ففاعليته يمكن أن يتم تقويمها عن طريق جمع البيانات في مناسبة أو أكثر قبل تطبيق البرنامج ومقارنتها ببيانات يتم جمعها بعد تطبيق البرنامج. وإذا ما كانت هناك إمكانية لتجريب البرنامج على شرائح أو مجموعة من المجتمع المستهدف فان ذلك أفضل من تجريبه على المجتمع بكامله، وبهذه الطريقة يستطيع المقومون عقد سلسلة من المقارنات بين المجموعة الواحدة بمرور الوقت وبين المجموعات المختلفة.

ب- **مجتمع الدراسة والعينة**: بعد تحديد مجتمع الدراسة على المقوم أن يقرر فيما إذا كانت الدراسة ستشمل كل المجتمع المستهدف أم أنه سيأخذ عينة منه.

هناك وجهتا نظر في موضوع استهداف مجتمع الدراسة. الأولى سياسية أو نفسية، تتعلق بأن الناس يشعرون بالضيق إذا لم يتم إشراكهم. فعلى سبيل المثال إذا أردنا دراسة اتجاهات المعلمين نحو استخدام التقويم الذاتي لتحديد كفاءة مدرسة ما فانه من الأفضل أخذ عينة من المعلمين من اختصاصات مختلفة بدلا من الاقتصار على معلمي اختصاص واحد فقط، حيث تؤكد الخبرة على ضرورة الشمولية التي لها مردود ايجابي من حيث تحقيق رضا المعلمين الوظيفي عندما تتاح الفرصة لجميعهم للإدلاء بوجهات نظرهم في الدراسة. أما وجهة النظر الثانية فتشير الى انه في حالة أن العينة الصغيرة جداً يكون هناك خطر عدم استطاعتنا تحديد أثر البرنامج حتى وان كان البرنامج فاعلاً (خطأ من النوع الثاني Type II error) فأخذ عينة من عشرة أشخاص من الذين شملهم البرنامج وعشرة من الذين لم يشملهم البرنامج ربما لن يؤدي إلى فروق ذات دلالة إحصائية حتى وان كان البرنامج ناجحاً.

من جانب آخر هناك وجهة نظر تدعم استخدام العينة العشوائية بدلاً من استهداف مجتمع الدراسة بكامله بسبب التكلفة المالية والوقت. اذ ان تقويم البرنامج يجب أن يتم وفق ميزانية محددة وفي وقت محدد. سواء اختار المقوم أن يشمل كل مجتمع الدراسة أم أنه اختار عينة، فلا يمكن تعميم النتائج

على مجتمع الدراسة بثقة تامة. ولذلك ففي مجتمع تعداده 1000 يستطيع المقوم في العادة التوصل إلى استنتاجات مفيدة إذا ما أخذ عينة قوامها 10% وحصل على نسبة استرجاع قدرها 75% من أن يحاول أن يشمل كل فرد في مجتمع الدراسة ويحصل على نسبة استرجاع 30%.

كما أن خصائص المجموعة المشاركة في الدراسة لها تأثيرها في معدل التسرب من الدراسة، فعلى سبيل المثال يصعب الحصول على معدل استجابة عالي على استبيان حول الرضا الوظيفي من قبل عمال النظافة في مؤسسة ما عنه من معلمي منطقة معينة. وهذا مرده أن معظم عمال النظافة من ذوي مستوى القراءة المتدني. لذلك يجب أن تكون أداة الدراسة قصيرة وبسيطة قدر الأمكان من أجل الحصول على معدل استجابة 50% أو أكثر.

ج- **المجموعات الضابطة والتجريبية**: عندما يتم توزيع المشاركين في الدراسة بشكل عشوائي على مجموعة لا تحصل على خدمات البرنامج الخاضع للتقويم فأنها تمثل مجموعة ضابطة حقيقية. وعندما يتم مقارنة المستفيدين من البرنامج بمجموعة غير مستفيدة من البرنامج ولكن لم يتم توزيعهم بشكل عشوائي ففي هذه الحالة يطلق على المجموعة (المجموعة المقارنة Comparison Group) فإذا ما كان الهدف من عملية التقويم معرفة ما إذا كان التغيير في المشاركين سببه البرنامج، فهنا يجب توزيع المشاركين إلى المجموعات التجريبية والضابطة بشكل عشوائي. وأن كان التوصل إلى علاقة السبب والنتيجة عادة ليس هدفاً في دراسات التقويم.

إن التوزيع العشوائي للمشاركين في العادة غير ممكن. ولكن هناك عدة طرق يمكن من خلالها الوصول إلى الهدف نفسه حتى وإن لم يكن للمقوم حكم تام في الموقف. ومنها على سبيل المثال لا الحصر مقارنة البيانات الأولية والديموغرافية للمشاركين في المجموعتين للتأكد من مدى التشابه بين أعضائها قبل البدء في الدراسة. كما أن إحدى الطرق المستخدمة لإعطاء ثقل أكبر لتفسير النتائج عندما يصعب الحصول على مجموعة ضابطة حقيقية هو تشكيل

أكثر من مجموعة مقارنة، ولذلك نقارن المجموعة التي خضعت للبرنامج محل التقويم بمجموعتين أو أكثر لم يشارك أعضاؤها في البرنامج مما يسهل من معرفة أثر البرنامج.

ولربما يتساءل البعض هل من الضروري وجود مجموعات مقارنة ليصبح التقويم ذو معنى، ألا توجد معايير مطلقة للأداء أو الجودة يمكن مقارنة النتائج بها؟

فعلى سبيل المثال إذا ما أدخلت مدرسة معينة نظاماً إلكترونياً لمعالجة بيانات الطلاب لتخفيف العمل على المعلمين ووجدنا بأن 80% منهم راضين عن البرنامج فهل في هذه الحالة نحتاج إلى مجموعة أو مجموعات مقارنة؟ طبعاً لا تبدو حاجة لذلك إذا ما وضعت إدارة المدرسة 80% كمعيار للرضا هدفاً. ولكن مع ذلك لعله من المناسب معرفة مستوى رضى المعلمين قبل إدخال النظام.

د- **اختيار المقاييس**: ان حجر الزاوية في القياس هو استخدام مقاييس مختلفة من مصادر مختلفة، فأهم مصدر للبيانات في برنامج التقويم هو العميل. ولكن يمكن لهذا العميل الإدلاء ببياناته بأكثر من طريقة. باستخدام التقرير الذاتي من خلال الإجابة على أسئلة مباشرة حول إنطباعاته ومستوى رضاه عن البرنامج. كما يمكن أن يجيب على مقياس الاتجاهات التي يمكن أن تتأثر بالبرنامج.

كما أن السلوك العملي يمكن أن يكون أحد مصادر المعلومات، فعلى سبيل المثال مستوى الرضا عن برنامج تدريبي لوظيفة ما يمكن قياسه عن طريق نسبة المنسحبين منه مقارنة بالبرامج الأخرى.

وهناك مصدر مهم آخر للمعلومات والبيانات يتمثل بأولئك الناس ذوي الصلة بالعميل كالأزواج والأقرباء والأصدقاء. ففي بعض الأحيان انطباعاتهم تكون أقل تحيزاً من انطباعات العميل نفسه. وأخيراً وليس آخراً إن العاملين في البرنامج يمكن أن يكونوا مصدرا مهما للمعلومات، ففي غالب الأحيان يكونوا مهنيين مهرة وانطباعاتهم عن البرنامج يمكن أن تكون لها مصداقية

أفضل سواء من العميل أو الفاعلين الآخرين. ولكن في الوقت نفسه انهماكهم في البرنامج يمكن أن يكون مصدراً للتحيز.

بجانب الحصول على ملعومات من مصادر مختلفة على المقوم أن يعرف أن استخدام مقاييس متعددة لكل مصدر ضرورة أو أمراً مرغوباً فيه. فعلى سبيل المثال بدلاً من استخدام مقياس واحد لقياس القلق لعله من المناسب استخدام مقياس يعطي معلومات عن الاكتئاب أو الإرهاق أو الإرتباك (فالحالة المزاجية مرتبطة بالقلق).

هـ- **جمع البيانات**: من سيتولى عملية جمع البيانات؟ يتطلب ذلك عادة وجود منسق لجمع البيانات في الموقع مهمته تحديد عناوين ومواقع عملاء البرنامج والعاملين فيه وغيرهم من مصادر البيانات بحيث يسهل التواصل معهم من قبل المقومين. وهذه مهمة صعبة وتحتاج إلى شخص موثوق به للقيام بها.

إن عملية جمع البيانات يجب أن تتسم بالمصداقية. فالمعلومات السرية يجب أن تعامل بسرية. والمعلومات التي يتم الحصول عليها على أساس من السرية لا يمكن الإدلاء بها قبل الحصول على موافقة صريحة بذلك. كما انه يجب احترام التزامات الأشخاص بالحفاظ على سرية المعلومات والمقومون ليس لديهم الحق المطلق في الحصول على كل المعلومات حتى إن كان عدم الحصول عليها سيعيق تقويم البرنامج. وهنا يجب التحقق من أن عملية جمع البيانات تتم بطريقة لا يتم فيها التجاوز على خصوصية العملاء.

و- **تحديد الإحصاء المناسب**: في أي عملية تقويم تبقى الحاجة قائمة في اختيار الإحصاء المناسب الذي يظهر مستوى الدلالة أو الفروق بين المجموعات وكذلك حجم الأثر (Magnitude of Effect) ويفضل دائماً استخدام عمليات إحصائية سهلة لأن النتائج سوف يتم تقديمها لمولي البرناج أو العاملين فيه والذين عادة ليس لديهم الخبرة الكافية في الإحصاء. فغير المختصين بالتقويم يجب أن يكونوا قادرين على فهم تفسيرات النتائج، لا

أن يكونوا منبهرين بمستوى الإحصاء المستخدم. ففي بعض الأحيان قد نرى إن (تحليل الانحدار المتعدد وتحليل الانحدار المصاحب) هما الأنسب، ولكن يجب الأخذ في الاعتبار إن مستخدمي نتائج التقويم يميلون إلى عمليات إحصائية اقل تعقيداً وللتغلب على ذلك يمكن استخدام العمليات الإحصائية الضرورية والتي تبين اثر البرنامج مع توضيح النتائج التي يتم التوصل إليها عن طريق النسب المئوية أو حتى الرسومات البيانية.

ز- **التقرير النهائي**: وأخيراً، يجب علينا في مرحلة التخطيط التفكير في شكل التقرير النهائي. فالأشكال والرسومات البيانية عادة ما تكون أكثر تفضيلاً من الجداول الإحصائية. وعندما تكون الجداول الإحصائية لا بد منها ينصح بالحد منها قدر الامكان ووضعها في ملحق في آخر التقرير. فشكل التقرير عادة ما يحدد نوعية البيانات والتحليلات الإحصائية المطلوبة.

الخطوة السادسة: تقديم خطة التقويم

بعد مراجعة أدبيات البحث والتفكير مليا في خطوات المنهجية التي سبق ذكرها في الخطوة الخامسة أعلاه يستطيع الباحث إعداد مقترح مكتوب لخطة التقويم ليتم تقديمه للعاملين في البرنامج. والهدف من ذلك التحقق من أن المقومين والعاملين في البرنامج متفقين على طبيعة وأهداف عملية التقويم ونوعية التقويم المطلوبة والمقاييس الإجرائية لأهداف البرنامج وجاهزية البرنامج للتقويم. فمن الأهمية بمكان من الناحية النفسية أن يكون العاملون في البرنامج على دراية كاملة بعملية التقويم، ويشعرون بالارتياح اتجاهها ولربما متحمسين لها. ولعل من المناسب مراجعة القضايا التي سبق مناقشتها خلال الاجتماعات التمهيدية مرة أخرى. فالعاملون في البرنامج قد يشعرون ببعض المشاكل الإجرائية التي لم تكن في الحسبان آنذاك. كما يمكن تأجيل جمع البيانات إذا ما شعروا بان البرنامج من النواحي الإجرائية غير مهيأ لعملية التقويم أو أن هناك عوامل خارجية قد تؤثر على عملية تفسير النتائج.

كتابة التقرير النهائي للتقويم Evaluation Report

مما لا شك فيه؛ أن طريقة كتابة وتقديم نتائج التقويم للمستفيد أو صانع القرار لها (أكبر الأثر في فهم التقويم، واستخدام نتائجه سواء لإجراء التعديل أو التطوير أو إيقاف البرنامج المقوَّم، ويؤكد هذه الحقيقة العديد من الباحثين ومنهم ليفتون وهيوز وغيرهم .(Leviton & Hughes, 1981)

وهناك العديد من النصائح يمكن أن تقدم للباحث للاهتمام بها عند كتابة تقرير التقويم ومنها:

1- أن تكون النتائج مكتوبة بأسلوب واضح ومفهوم يستطيع صانع القرار فهمها واستيعابها بسهولة.

2- أن تكتب النتائج بأسلوب إجرائي يوضح أساليب تطبيقها لمعالجة الحالات التي تم تقويمها.

3- ضرورة الابتعاد ?كلما أمكن- عن خلط النتائج بالوسائل الإحصائية أو المعادلات الرياضية، مما يعقد الأمر على المسؤول لفهمها لاسيما إذا لم تكن لديه الخبرة الكافية للتعامل مع الأساليب الإحصائية، مما يؤدي إلى ضياع الاستفادة منها في التطبيق، ولا بأس في أن يخصص جزءاً منفرداً للأمور الإحصائية في تقرير التقويم.

4- يفترض بالمقوم في كتابة ملاحظاته أو استنتاجاته أن لا يستخدم الأسلوب القاطع بل أن يعترف بنسبية هذه النتائج واعتمادها على الظروف والمتغيرات المحيطة بالموضوع المقوَّم، كما يجب الحذر من استخدام أسلوب التعميم؛ لأن موضوعات التقويم كثيراً ما تكون متفردة، إلا إذا كانت هناك حجج وأسانيد علمية واضحة تدعم إمكانية التعميم.

5- يفترض بالمقوم أن يوضح في تقريره الأهمية العلمية المتوخاة من تطبيق مقترحاته، وأهميتها في تحسين البرنامج المقوَّم، كما أنّ عليه أن يوضح أنّ نتائج بحثه تتمتع بدرجة كافية من الصدق لموضوعية الأسلوب المستخدم أو التصميم المستخدم، وسلامة المقاييس والأدوات المستخدمة في توفير البيانات لما لذلك من أهمية في إقناع المسؤولين للأخذ بها وتطبيقها في الميدان.

وقد توصل كل من (وايز وبوكوفالس Weiss & Bucuvalas 1980) من خلال دراستهما عن ظروف استخدام نتائج التقويم إلى أنّ متخذي القرار يستخدمون معيارين للتمييز بين التقويمات التي يريدون استخدام نتائجها، وهما معياري الفائدة Usefulness والصدق Validity، حيث يستند معيار الفائدة على أساسين هما: نوعية أسلوب التقويم، ومدى انسجام نتائجه مع نتائج البحوث السابقة وتوقعات العاملين في نفس المجال.

أما معيار الصدق فيُستخدم للتأكد من مصداقية النتائج المستندة على مصداقية أدوات التقويم، ومع ذلك فقد أشار (روسي وفريمان Rossi & Freeman 1982) إلى أنّ موضوع استخدام نتائج البحوث التقويمية عملية معقدة تتداخل فيها متغيرات عديدة ومتنوعة.

6- إنّ تقرير التقويم يجب أن يكون شاملا لكل عناصر الموضوع المقوَّم ابتداءً من الأهداف العامة والجزئية وحتى تقويم النتائج مرورا بالعمليات، حيث يفترض بالمقوم أن يوضح كل شيء، ويجمع بيانات ومعلومات واسعة عن العناصر المكونة للبرنامج؛ من أجل الاستفادة منها في مجالين:
أ. توضيح صورة البرنامج أمامه، وأمام متخذ القرار.
ب. تحديد علاقة تلك العناصر بنتائج البرنامج تحت التقويم.

كما أن على الباحث أن يقف عند عنصر العمليات ليدرسها بعمق ودقة، لأن العمليات التي تتضمن التصميم والمدخلات وتفاعلها تؤثر إمّا سلبا أو إيجابا على نوعية النتائج.

7- يفترض بالمقوّم أن يلاحظ التأثيرات الخاصة بالبرنامج المقوَّم على المحيط الذي يعمل فيه، وأن يركز أيضا على مدى إيفاء البرنامج لحاجات المستفيدين منه، وعلى الباحث أنْ لا ينسى الاستئناس بآراء ومقترحات العاملين في ذلك البرنامج إذْ أنّ لبعضهم الخبرة الكافية لتحديد جوانب الضعف والقوة في البرنامج المقوم، وعلى المقوِّم أن لا يهمل أية ملاحظة يسمعها وإنما يمحصها بدقة ليرى إن كانت من الفاعلية بحيث تستحق ذكرها في التقويم.

8- ومن الأمور الضرورية في تقرير التقويم توضيح المعايير التي استند عليها في إصدار أحكامه، وأن يوضح أن تلك المعايير معروفة ومجربة، ويفضّل أن يكون لها طابع كمّي لتسهل مقارنة الواقع بها، والخروج باستنتاجات موضوعية تؤدي إلى أحكام موضوعية.
9- على الباحث أن يوثق علاقته بصانع القرار، وأن يستشيره في استيضاح الكثير من الأمور ويشجعه على التعاون معه، ويتبنى أفكاره واقتراحاته الإيجابية؛ لأن ذلك سيؤدي إلى الشعور بأن له دورا في عملية التقويم ممّا يجعله أكثر تحمسا لتطبيق نتائج عملية التقويم.

توصيات لزيادة استخدام نتائج التقويم Utilization of Evaluation Results

يقترح كل من سولومون وشورتل (Solomon & Shortell, 1981) مجموعة من التوصيات التي يمكن أن تؤدي إلى زيادة استخدام نتائج التقويم وهي:
1- على المقوم أن يفهم الإطار الفكري / المعرفي لصاحب القرار؛ كيف يفكر؟ ما قيمه؟ ما آراؤه؟ لكي يلائم بين تقريره وتلك الأطر.
2- يجب أن توقت نتائج البحث بتقرير التقويم لتقدم في الوقت المناسب عند الطلب، إذْ أنّ هناك أمورا سلبية تنشأ عن تأخر تقرير التقويم، أو تقديمه بعد أن تكون الحاجة له انتفت أو مرت زمنيا.
3- يجب أن تحترم التقويمات تعليقات صانع القرار، وأن تهتم بأخذها بنظر الاعتبار ولاسيما إذا لم تكن متعارضة مع الأسس الخاصة بعملية التقويم، ويجب أن لا يكون هناك تعارض جوهري بين تلك التعليقات وأفكار وإجراءات المقوّم.
4- يجب أن تكون أسس استخدام نتائج التقويم موضحة، وكجزء مهم من تقرير التقويم، وأن توضّح بشكل إجرائي قابل للتطبيق مع توضيحات تفصيلية تساعد صانع القرار على استخدامها في التعديل أو التغييرات المطلوبة على البرنامج.
5- يجب أن تتضمن التقويمات تقييما لعملية الاستخدام، فعلى المقومين

وصانعي القرار أن يتفقا على معايير لاستخدام نتائج التقويم، إلى جانب اشتراكهم في فهم الأغراض التي أجريت عملية التقويم من أجلها، وفي هذا المجال يجب أن تتماشى هذه المعايير مع توقعات استخدام تلك النتائج وفق توقعات زمنية أو مهنية، أي كيفية توقيت التعديلات المرغوب إجراؤها في البرنامج المقوم، وما هي أولويات التغيير؟ وكيفية تطبيقه؟ وغير ذلك.

العلاقة بين المقوم ومتخذ القرار

يؤكد بعض الباحثين ومنهم جوبا وستفلبيم (Guba & Stufflebeam 1988) على أن الدور الأساسي للمقوم هو أن يقدم معلومات وافية لمتخذ القرار. لذلك يجب أن يكون المقوم متخصصاً في عمليات تشخيص وجمع وتحليل وكتابة تقرير جيد عن عملية التقويم. وإن عمل الأحكام المختصرة يجب أن تترك لمديري البرنامج، بينما إقترحَ ستيك (Stake 1967) أن المقوم هو الذي يجب أن يعمل الأحكام عن البرنامج. ومن وجهة نظره واستنتاجاً من آرائه؛ ولذلك فإن مخرجات الدراسة التقويمية يجب أن تتضمن:

أ. وصفاً كاملاً لجميع المتغيرات المهمة التي تؤثر على سير البرنامج.

ب. مجموعة من الأحكام الواضحة عن البرنامج.

فإذا ما اكتملت هذه الأمور فإن أي شخص متخصص سيصل إلى نفس النتائج والاستنتاجات حول الموضوع المقوم.

باحثون آخرون أكدوا على أن المقوم يجب أن يعطى مسؤولية صنع قرارات مستندة إلى نتائج الدراسة التقويمية. مضيفين إلى أن المنطق وراء هذه المسؤولية وهو أن المقوم هو الشخص الوحيد الذي يكون عارفاً بكل الأوجه المهمة في البرنامج. لاسيما أنه قام بجمع البيانات اللازمة عن البرنامج فأصبح بوضع يستطيع معه الحكم على جميع أوجه البرنامج، بينما يكون صانع القرار هو من يُعْلَم بهذه النتائج والأحكام. وإن قيام صانعي القرار بعملية إصدار

الأحكام - كما أشار لها البعض- عملية متحيزة. لذلك يجب أن تعطى مسؤولية إصدار الأحكام إلى المقوم نفسه.

ويؤكد باحثون آخرون أن من المثمر والمنطق أن يتعاون كل من المقوم والمسؤول عن البرنامج لاتخاذ القرارات النهائية حول البرنامج المقوم. لأنهما سبق أن عرفا وراجعا كل أوجه البرنامج معاً، لذلك يمكن لهما أن يراجعا وينقحا تقرير التقويم قبل كتابته بشكله النهائي للتوصل إلى عمل الأحكام القيمة معاً.

إن المقوم يجب أن يقدم تقريره النهائي والأحكام والتوصيات لمتخذ القرار حيث يراجع المسؤول التقرير ليوافق أو يعارض الأحكام النهائية التي قد يعتقد أنها غير مستندة إلى بيانات واضحة. فإذا حدث نوع من عدم الاتفاق فيمكن إعادة جمع بيانات أخرى عن الموضوع لوضع أحكام نهائية. وعند ذلك يشعر الرأي العام للمستفيدين من البرنامج كيفية توصل الطرفين إلى الأحكام النهائية. والوثوق بأنهما كانا أهلاً لتحمل المسؤولية التي تؤهلهم لإصدار الأحكام (Worthen & Sanders, 1987).

معايير ومحكات التقويم التربوي

سبق وأشرنا في صفحات سابقة من الكتاب إلى أن عملية التقويم يجب أن تنتهي بإصدار أحكام قيمية عن الشيء المقوّم أي مستوى جودته أو نوعيته. ولأجل أن تتصف تلك الأحكام بالموضوعية لابد وأن تستند إلى محكات أو مستويات ضمنية أو صريحة تتسم بالصدق والموضوعية أيضاً. وتتمثل تلك المحكات أو المعايير في صور متنوعة كأن تكون بصورة الفائدة التي يحققها برنامج ما مثال ذلك عندما نقوّم جودة برامج تربوية في ضوء ما تسهم به في تنمية مهارات مطلوبة، أو تحقيق احتياجات فئة مستهدفة، أو تقويم سياسات معينة في ضوء ما يؤدي إلى تطور نوعي أو كمي لأفراد المجتمع، وقد يكون الحكم على البرنامج من خلال مدى ما تحقق من الأهداف التي يسعى البرنامج لتحقيقها.

هذه الأهداف والفوائد تعتبر محكات Criteria أو مستويات Standards يتم

التقويم في ضوئها، وكلما كانت هذه المحكات واضحة ومحددة ولها صور كمية كلما يسهل على المقوّم مقارنة واقع البرنامج بها وإصدار الحكم على ذلك البرنامج، لكن الأهداف على سبيل المثال قد تكون غامضة أي غير محددة ومتداخلة مع غيرها من الإجراءات والمؤشرات الخاصة بالبرنامج مما يجعل من الصعب إجراء المقارنة وإصدار الأحكام، لذلك ينبغي أن يعرف المقوّم من صانع القرار ما إذا كانت هناك معايير تحكم العملية، وأن يتفقا على محكات واضحة لأجل الاستناد إليها في قبول أو استبعاد معلومات وبيانات معينة. ويؤكد بعض الباحثين في مجال التقويم إن محكات التقويم لا تقتصر على الجوانب الفنية أو المنهجية، بل تتطلب النظر إلى مكونات ثلاث لعملية صنع القرارات، حيث أن كلاً من هذه المكونات تتطلب محكات فاعلية مختلفة، ويرى سشمان (Suchman, 1967) أن هذه المكونات هي:

1- مكونات معلوماتية Information Component

وتتمثل في توفر نظام مرجعي، أو نموذج عملي تستند عليه في انتقاء المعلومات وجمعها وتصنيفها وهي تمثل محك موائمة الفعل.

2- مكونات تفسيرية Interpretive Component

وتتضمن اختيار نظام تفضيل أو إطار قيمي يمثل الوسائل والغايات المرغوبة أو المناسبة لصنع القرارات، وهي تمثل عملية ترجيح وزني، أو قاعدة تستند إليها في صنع القرار الخاص بتحديد الأولويات من بدائل متعددة.

3- مكونات إجرائية Action Component

وتتمثل مبدأ قبول المخاطرة أو المجازفة في حل المشكلات، حيث تكون وظيفة المحكات تقليل الخطأ، وفي هذه العملية يتم تقدير الكلفة المترتبة على الخطأ مقابل النواتج المطلوبة.

ومن الجدير بالذكر أن التقويم واستخدام هذه المحكات ليس عملاً يتم مرة واحدة خلال عمر البرنامج بل يمثل سلسلة مستمرة من تحديد الاختلافات وإيجاد حلول لها أثناء مسيرة البرنامج، وبذلك لا تكون المحكات جامدة

(ثابتة) وآلية، بل إنها دائمة التغيير تتطور حسب ظروف البرنامج من جهة ونظرة المقوّم وفلسفته ونظرة متخذ القرار أيضاً.

ولا تقتصر أهمية تحديد المعايير والمحكات الخاصة لغرض إصدار الأحكام القيمية الناشئة عن عملية التقويم، بل إنها تمتد إلى تحديد مواصفات وشروط إجراء عملية التقويم والعناصر المتعلقة بها وبضمنها الباحث نفسه. لقد سارعت جمعيات ومنظمات التقويم العالمية بإصدار معايير خاصة بها تمثل شروطاً موضوعية ومواصفات محددة تطلب توفرها بعملية التقويم وفي القائم بالعملية (المقوّم) حيث يقترح ستفلبيم Stufflebeam نوعين من المعايير ضرورية لعملية التقويم هما:

أولاً: معايير رابطة التقويم التربوي الأمريكية

Joint Committee Standards for Educational Evaluation (JCSEE, 1981)

وتتكون من (30) معيار طورتها تلك الرابطة عام 1981 نقدم فيما يأتي ملخصاً لها:

أ. معايير المنفعة:

وتهدف إلى التأكد من أن التقويم سيخدم احتياجات من يطلب القيام به، ومنها:

1- تحديد المستفيدين من التقويم أو المتأثرين به لغرض خدمة احتياجاتهم.
2- مصداقية المقوّم ومدى تمكنه من إجراء التقويم حتى نضمن الثقة في نتائج تقويمه وقبول مصداقيتها.
3- جمع المعلومات واختيارها أي تُجمع وتختار لكي تجيب على أسئلة محددة حول موضوع التقويم وتستجيب لحاجات المستفيدين من البرنامج ومن التقويم.
4- التفسير القيمي: وصف الإمكانات والإجراءات والمبررات المستخدمة في تفسير النتائج التي وصل إليها التقويم.
5- وضوح التقارير: يجب أن يصف التقرير موضوع التقويم، محتواه،

غرضه، إجراءاته ونتائجه كي يتمكن المستفيدون فهم كيف وضع، ولَمَ وضع، وما نوع المعلومات التي تم الحصول عليها، والاستنتاجات التي تم التوصل إليها والتوصيات.

6- نشر التقرير: أي وصوله للمستفيدين الذين يحق لهم معرفته حتى يتمكنوا من الحكم على التقويم والاستفادة من نتائجه.

7- توقيت التقرير: أي يجب أن يظهر التقرير في وقت يسمح بالاستفادة من نتائجه والمعلومات التي يحتويها.

8- تأثير التقويم: أي أن يخطط التقويم وينفذ على نحو يشجع متابعته من قبل المستفيدين.

ب- معايير الجدوى:

وتهدف هذه المعايير إلى التأكد من أن التقويم واقعي، متعقل، دبلوماسي، ومنها:

1- الإجراءات العملية: أي أن تكون إجراءات التقويم عملية وتحصل على المعلومات بدون إرباك للعمل.

2- المقبولية: يجب أن يخطط التقويم وينفذ مع الأخذ في الاعتبار قبول المجموعات التي لها مصلحة في البرنامج من أجل أن يُضمن تعاونهم ومنع محاولاتهم لتغيير نتائج التقويم أو تزييفها.

3- فعالية التكلفة: أي أن يوفر التقويم معلومات كافية تبرر تكلفة جمعها.

جـ. المعايير القانونية:

ويقصد بها التأكد من أن التقويم في تنفيذه يجب أن يراعى الجوانب القانونية والأخلاقية مع الأخذ بنظر الاعتبار مصلحة المشاركين فيه والمتأثرين بنتائجه.

1- الالتزام الرسمي: يجب أن تحدد التزام الأطراف المتفقة على إجراء التقويم (ماذا سينفذ، كيف ينفذ، من ينفذ، متى يُنفذ) ويتم إبرام عقد كتابي يلتزم به الأطراف المتعاقدة بجميع الشروط أو بعضها.

2- تعارض المصلحة: عندما يظهر تعارض في المصلحة بين المسؤول والقائم في

التقويم، يجب أن يحل بوضوح وصراحة حتى لا يؤثر على التقويم ونتائجه.

3- المكاشفة الكاملة الصريحة: أي أن تكون تقارير التقويم المكتوبة والشفوية، مباشرة وصريحة فيما يتعلق بالنتائج ومحددات التقويم.

4- حق الجمهور في الإطلاع: يجب أن تحترم الأطراف الرسمية في التقويم وتؤكد على حق الناس في الإطلاع على التقويم ضمن الحدود التي تسمح بها المبادئ والحقوق الأخرى لا سيما التي تتعلق بسلامة الناس والحفاظ على خصوصياتهم.

5- حقوق الأفراد: يجب أن يضمن تصميم التقويم وتنفيذه صيانة حقوق الأفراد ومصالحهم.

6- التفاعل الإنساني: أن يحترم المقوّم كرامة الأفراد وقيمهم عند تفاعلهم مع القائمين على التقويم.

7- التقارير المتوازنة: أن يكون تقرير التقويم موازناً ومنصفاً في تشخيصه لنقاط القوة أو الضعف، حتى يتم البناء على جوانب القوة ويتم معالجة جوانب الضعف.

د- **معايير الدقة:**

تهدف هذه المعايير إلى التأكد من أن التقويم سيكشف ويقدم معلومات كافية عن البرنامج المقوم وتقرر قيمته وجدارته.

1- تحديد الموضوع: وذلك من خلال فحص موضوع التقويم فحصاً كافياً ليتم تحديده بوضوح سواء كان برنامجاً، مشروعاً، أو مادة تعليمية.

2- تحليل المحتوى: فحص محتوى موضوع التقويم بشكل مفصل لتحديد احتمالات تأثيره على الموضوع.

3- الأغراض والإجراءات: مراقبة أغراض التقويم وإجراءاته بشكل مفصل بحيث يتم تحديدها وتقويمها.

4- مصادر المعلومات: يجب وصف مصادر المعلومات بشكل مفصل بحيث يمكن تقويم مدى كفايتها للتقويم.

5- طرق القياس: يجب اختيار أدوات القياس وجمع المعلومات وتطبيقها بالطريقة التي تضمن صدق التفسيرات التي يتم التوصل إليها.
6- ثبات القياس: يجب اختيار أدوات القياس وجمع المعلومات وتطبيقها بالطريقة التي تضمن أن المعلومات ثابتة ودقيقة.
7- الضبط المنظم للبيانات: يجب مراجعة البيانات المستخدمة في التقويم وتنقيحها حتى لا تكون نتائج التقويم خاطئة.
8- تحليل المعلومات الكمية: يجب تحليل المعلومات الكمية بشكل مناسب ومنظم لتضمن توافقها مع التفسيرات.
9- تحليل المعلومات النوعية: يجب تحليل المعلومات النوعية بشكل مناسب ومنظم لتضمن توافقها مع التفسيرات.
10- وضوح الخاتمة: يجب أن تكون خاتمة التقويم واضحة حتى يمكن تقويمها.

ثانياً: معايير رابطة بحوث التقويم: Evaluation Research Society

وضعت هذه المعايير عام 1994 وهي مكونة من أربعة معايير يفترض بالمقوّم أن يلتزم بها وهي:
1- أن يفيد التقويم في تلبية احتياجات المسؤولين عن تطبيق نتائجه مجيباً على الأسئلة التي تعينهم، كما يجب أن تقدم تقارير التقويم في الوقت المحدد بحيث توفر تغذية راجعة للبرنامج تكون مفيدة في تحسينه وتطويره.
2- يجب أن يكون التقويم مجدياً ويستخدم الإجراءات التي من شأنها أن لا تؤدي إلى إرباك البرنامج من خلال تطبيقها، ويأخذ الاحتياجات الكافية لمنع أي تلاعب بأغراض التقويم وأن ينفذ التقويم بمنتهى الفاعلية.
3- يجب أن يكون التقويم أخلاقياً، مبنياً على إتفاق واضح يحدد المسؤوليات، وأن لا تخرق النتائج لتناسب مصالح أفراد او مجموعات معينة بما يؤدي إلى ضياع حقوق الآخرين المستفيدين من البرنامج.
4- يجب أن يكون التقويم دقيقاً في وصفه للموضوع المقوّم، بحيث يوضح

نقاط القوة والضعف سواء في خطة التقويم وإجراءاته، وخاتمته وأن يبتعد عن التحيز، وتكون نتائجه صادقة وثابتة (الدوسري، 2000). ولزيادة الفائدة يضيف الباحثون نوعاً آخراً من المعايير هي:

ثالثاً: معايير الجمعية الأمريكية للتقويم

American Evaluation Association (AEA)

وضعت جمعية التقويم الأمريكية عام 1975 مجموعة من المعايير الخاصة بالتقويم وإجراءاته، إرتأى الباحثان عرض أهم تلك المعايير بشكل مختصر وكالآتي:

أ. معايير الاستخدام:

وتضمن هذه المعايير قيام المقوّم بتلبية حاجات ورغبات مستخدم المعلومات الناشئة عن التقويم من خلال:

1- تحديد المستفيد: أي تحديد الأشخاص المشمولين أو المتأثرين بعملية التقويم حتى يمكن إشباع حاجاتهم.
2- مصداقية المقوّم: يجب أن يكون الشخص القائم بعملية التقويم موثوقاً به وجديراً بالقيام بالتقويم، من أجل أن تحوز نتائج التقويم على أعلى درجة من المصداقية والقبول.
3- تحديد القيم: إن التطورات والأساليب والمنطق المستخدم في تفسير نتائج التقويم يجب أن توصف بعناية حتى تكون الأحكام القيمية مبنية على أسس واضحة.
4- وضوح التقرير: يجب أن تتضمن تقارير التقويم وصفاً دقيقاً للبرنامج المقوّم سواء من حيث بيئته، عملياته، أساليبه، أو نتائجه وبذلك ستكون المعلومات الأساسية الخاصة بالبرنامج مفهومه بسهولة.

ب. المعايير المعقولة:

وذلك من أجل ضمان أن يكون التقويم واقعياً معقولاً، دبلوماسياً واقتصادياً ومنها:

1- الأساليب العملية: يحب أن تكون أساليب التقويم عملية حتى تكون الفوضى بأقل ما يمكن أثناء عملية جمع المعلومات.
2- فاعلية التكلفة: يجب أن يكون التقويم فعالاً ومنتجاً لمعلومات ذات قيمة عالية، حتى يمكن تبرير تكلفته العالية.

جـ- المعايير المناسبة:

تهدف هذه المعايير ضمان شرعية التقويم وأخلاقيته، ويهدف لخير المساهمين فيه، أولئك المتأثرين بنتائجه، ومنها:

1- توجيه الخدمات: يجب أن تصمم التقويمات لتساعد المنظمات على أن تلبي وتخدم حاجات مدى واسعاً من المشاركين والمستهدفين بفاعلية.
2- حقوق القوى البشرية: يجب أن تصمم التقويمات وتدار بحيث تحترم وتحمي حقوق ورفاهية العناصر البشرية في البرنامج.
3- المسؤولية المادية: أن مخصصات المقومين ومصاريفهم المالية يجب أن تعكس أساليب المحاسبة الواضحة، والمسؤولية الأخلاقية، لذلك فإن المصروفات يجب أن تحسب بطريقة هادفة ومناسبة.

د- معايير الدقة:

وتضمن هذه المعايير أن التقويم سوف يكشف عن معلومات تقنية مقنعة ووافية عن خصائص وتقرير قيمة أو جدارة البرنامج قيد التقويم ، ومن هذه المعايير:

1- توثيق البرنامج: يجب أن يوصف البرنامج المقوّم ويوثق بوضوح ودقة، وبذلك يكون قد تم تشخيصه بوضوح.
2- تحليل البيئة: يجب أن تفحص البيئة التي يوجد فيها البرنامج بشكل مفصل وكاف، وذلك لتشخيص العوامل التي تؤثر على البرنامج بشكل واضح ودقيق.
3- وصف أغراض البرنامج وأساليبه بشكل واضح.

4- صدق المعلومات: أي أن الأساليب المستخدمة لجمع البيانات عن البرنامج يجب أن تتمتع بالصدق الكافي لتكون المعلومات نفسها صادقة.
5- ثبات المعلومات: أي أن الأساليب المستخدمة في جمع البيانات عن البرنامج يجب أن تكون ذات ثبات كاف لأجل أن تأتي البيانات الناشئة عنها تتمتع بالثبات المطلوب.
6- تحليل المعلومات الكمية: المعلومات الكمية في التقويم يجب أن تحلل بأسلوب منظم ومناسب لأجل أن تتم إجابة أسئلة التقويم بشكل فاعل.
7- تحليل المعلومات النوعية: يجب أن يتم تحليل المعلومات النوعية بشكل منظم ومناسب يضمن توفير إجابة فاعلة لأسئلة التقويم.
8- الاستنتاجات المبررة: يجب تبرير الاستنتاجات التي يتم التوصل إليها في عملية التقويم بشكل واضح حتى يستطيع المسؤول عن البرنامج الانتفاع بها.
9- التقرير المتكامل: الأساليب المستخدمة في وضع تقرير التقويم يجب أن لا توجه بمشاعر شخصية أو تحيزات ضد أي طرف من التقويم، وأن يعكس التقرير نتائج التقويم بشكل موضوعي.
10- ما بعد التقويم: إن التقويم نفسه يجب أن يقوم بشكل بنائي وختامي في ضوء معايير دائميه، حتى تكون إدارته موجهة ومنجزة، ويتمكن المسؤولون عن البرنامج من فحص وتشخيص جوانب القوة والضعف. (Mac Nmara, 1998).

إدارة عملية التقويم

لعله من المهم قبل الشروع في عملية التقويم وضع خطة تفصيلية للمهام المختلفة التي تتضمنها العملية، كأن تكتب أسئلة التقويم المختلفة وبجانبها المهام اللازمة للإجابة على السؤال، والوقت اللازم لإنجاز كل مهمة، وتاريخ البدء بها والانتهاء منها، والأشخاص الذين سيقومون بإنجازها، والمستلزمات الضرورية لتنفيذها، وتكلفة إنجازها.

ولإدارة عملية التقويم بشكل فعال، ينصح باستخدام مخطط جانت Gantt Chart، والذي يتكون من خطين أفقي ورأسي، يتم كتابة خطوات أو مراحل عملية التقويم متسلسلة حسب بدء حدوثها في الخط الرأسي، والزمن المخصص لعملية التقويم مقسما إلى أسابيع في الخط الأفقي. ويتم رسم خط أفقي لكل مهمة يبدأ من تاريخ بدء العمل في المهمة وينتهي بنهاية الزمن المحدد لإنجازها، حيث يبين طول الخط الزمن المخصص لإنجاز كل مهمة. ومن خلال النظر إلى مخطط جانت يستطيع المقوم أو أي شخص آخر معرفة خطوات عملية التقويم المختلفة وموعد حدوثها والانتهاء منها، وما يميز (مخطط جانت) أنه سهل التحضير، ويساعد القائم بعملية التقويم على تتبع تقدمه في العملية، وإذا ما التزم المقوم بإنجاز المهام في موعدها فإنه يضمن الانتهاء من عملية التقويم في الوقت المخطط لها.

مراجع الفصل الرابع

1- الدوسري، إبراهيم بن مبارك (2000)، الإطار المرجعي للتقويم التربوي. الرياض: مكتب التربية العربي لدول الخليج.

2- Carey, R. G., & Posavac, E.J. (1978) **Program evaluation of a physical medicine and rehabilitation unit: A new approach. Archives of Physical Medicine and Rehabilitation**, 59; 330-333.

3- Cronbach, L.J. (1977). **Remarks to a new society Evaluation Research**. Society News letter 1: 1-3.

4- Ellsworth, R.B. (1979). **Evaluating treatment outcomes in mental health services.** In E. J. Posavac. Impacts of Program Evaluation on Mental Health Care. Boulder, Colorado: Eestview press.

5- Guba, E.G., & Stufflebeam, D.L. (1968). **Evaluation: The process of stimulating, aiding, and abetting insightful action**. Paper presented at the 2nd Phi Delta kappa symposium for Professors of Educational Research, Colorado.

6- Horst, P., et al. (1974). Program management and the federal evaluator. **Public Administration Review**, 54; 300-308.

7- Levition, L. C., & Hughes, E.E. (1979). **Vtilization of Evaluations**. Ilinois. Northwestren University Center for health service and Policy Research.

8- MacNmara. C. (1998). **Educational Evaluation Standerds of American Evaluation Association.**

9- Posavac, E.J., & Carey, R.G. (1985). **Program Evaluation: Methods and Case Studies**. New Jersey: Prentice-Hall, Inc.

10- Rossi, P.H., & Freman, H.E. (1982). **Evaluation: Asystematic Approach (2 nd ed).** California: SAG Publications, Inc.

11- Solomon, M. A., & shortell, S.M. (1981). Desining Health Policy Research for Utilization. **Health Policy Quarterly** 1: XX.

12- Stake, R.E. (1967). The Countenance of Educational Evaluation. **Teacher College Records**, 68, 523-540.

13- Suchman, E. (1967). **Evaluative Research**. New York: Russell Sage.

14- Weiss, C.H., & Bucuvalas M.J. (1980). Trust tests and Utility Tests: Decision - Makers' frames of Reference for social science Research. **American Sociological Review**, 45 (April): 302-313.

15- Worthen, B. R., & Sanders, J.R. (1973). **Educational Evaluation: Theory and Practice**. California: Wadsworth Publishing Company, Inc.

الفصل الخامس

أدوات التقويم

الملاحظة

المقابلة

قوائم المراجعة (الشطب)

الاثنوغرافيا

الاستبيان

5

أدوات التقويم

يتفق العاملون في مجالي البحث والتقويم على ضرورة تنوع الأدوات المستخدمة في عملية التقويم من أجل جمع أكبر قدر ممكن من المعلومات المتعلقة بالموضوع المراد تقويمه للتوصل إلى أحكام أكثر دقة وموضوعية عنه. إن تعدد أدوات التقويم ومن خلال المعلومات التي توفرها تزيد من وضوح الموقف المقوّم، ففي تقويم برنامج تربوي معين، لابد من استخدام الملاحظة لرصد عمل البرنامج وسير حركته بشكل عام، ومن خلال المقابلة مع القائمين على البرنامج أو المستفيدين منه يمكن الحصول على معلومات مفيدة عن أهداف البرنامج وطبيعتها وعن عملياته وتلمس مشاعر الرضا من عدمها عن ذلك البرنامج، وعندما يكون عدد العاملين فيه كبيراً يمكن استخدام الاستبيان لرسم صورة واضحة عنه، ونفس الشيء يمكن أن ينطبق على عملية تقويم الطالب، فالأعمال والنشاطات المطلوب ان يقدمها الطالب من تقارير، وعروض ومجسمات وامتحانات قصيرة في فترات معينة من الفصل الدراسي وملاحظات لسلوكه في الصف وخارجه تكوّن بالتأكيد صورة واضحة وجلية عن مستوى الطالب الدراسي وحتى الاجتماعي والنفسي.

وتصنف أدوات التقويم التي تستخدم لجمع المعلومات عن الموضوع المَقَوَّم تصنيفات شتى لعل أهمها تصنيف تلك الأدوات وفقاً لنوع التقويم بين التقويم الكمي والتقويم النوعي. ففي الاسلوب الكمي الذي يعتمد على البيانات الرقمية المستندة الى التحليل الاحصائي سواء في مرحلة بناء الأدوات أو في جمع البيانات وتحليلها تعتبر الملاحظة المقيدة والمقابلة المقيدة والاستبيان والاختبارات والمقاييس وتحليل المحتوى أهم تلك الأدوات.

أما الأدوات التي تستخدم في التقويم النوعي فأهمها الملاحظة الطبيعية

والمقابلة المفتوحة وتحليل الوثائق الخاصة بالأفراد والسجلات والبريد الالكتروني. وسيتم في الصفحات القادمة عرضاً مفصلاً لبعض تلك الأدوات ونقاط ضعفها وقوتها وأساليب استخدامها.

إن أهمية أدوات التقويم تتجلى في قدرة الباحث على تصميمها بالشكل الصحيح قبل استخدامها لجمع المعلومات عن الموضوع المراد تقويمه من خلال تأكده من توفر الخصائص السيكومتريه الضرورية لا سيما الصدق Valiolity والثبات Reliability والشمول Comprehensiveness والخصائص الأخرى المطلوبة.

الملاحظة: Observation

تعتبر الملاحظة واحدة من أهم أدوات البحث والتقويم، وقد استخدمها الإنسان منذ بدايات وجوده على الأرض لمتابعة وتقويم الظواهر الطبيعية المحيطة به ليقرر بعد ذلك الابتعاد عنها أو الاقتراب منها، وليمد الجسور معها عندما يتأكد من عدم ضررها له، ويسخرها لمصلحته.

ولا زالت الملاحظة حتى الآن تستخدم في كافة المجالات العلمية الصرفة والإنسانية، حيث يستفيد منها الباحثون لجمع المعلومات التي توظف لتحقيق أهداف البحث العلمي. وفي ميدان التقويم تستخدم الملاحظة في مجالات كثيرة في معظم أنماطه، سواء التقويم الكمي أو التقويم النوعي، وفي تقويم المؤسسات والبرامج التعليمية تلعب الملاحظة دوراً كبيراً في توفير البيانات اللازمة للتعرف على البرنامج من حيث وصفه أو متابعة حركة سيره وحتى تطوره خلال فترة من تاريخ تأسيسهُ. ويُشير الباحثون إلى أن الملاحظة تختزل التحيزات التي تكمن في استخدام الأساليب الإخرى والتي غالباً ما تتأثر بالتقرير الذاتي للفرد، إذ أنها (الملاحظة) تعكس الواقع دون تضخيم أو تزييف وتوفر بيانات أكثر دقة وواقعية.

وتتطلب الملاحظة من المقوِّم أن يحدد المتغيرات التي يُفترض أن تُلاحظ بدقة ومعرفة مسبقه لأهميتها، فعندما نريد تقويم تدريس معلم مادة اللغة

الإنجليزية في مرحلة تعليمية معينة، علينا أن نهتم بملاحظة المتغيرات الأساسية في تدريس الموضوع ، ثم نحدد المكان والزمان اللذان تجري فيهما عملية التدريس، ثم نضع تفصيلات وعناصر التدريس التي يجب أن نركز الإنتباه عليها بشكل دقيق من ذلك مثلاً: أسلوب عرض المادة على الطلاب، توجيه الأسئلة، تشجيع الطلاب على التفكير قبل الإجابة ، قدرة المعلم على تصحيح أخطاء الطلاب في وقت وقوعها، تلفّظ المعلم للكلمات بشكل صحيح، والواجبات الصفية والبيتية التي يكلف بها طلابه. . .الخ، ثم نقوم بتسجيل هذه الملاحظات في استمارات مقننه لغرض تكميمها وتفسيرها إذا كان تقويمنا كمياً. أما في التقويم النوعي فيحاول المقوِّم جمع أكبر قدر من الملاحظات عن الموقف بوضعه الطبيعي ويسجلها ثم يقوم بتلخيصها والاستفادة من خبراته ووجهات نظره في تفسير الملاحظات، أي إن المعلومات التي يجمعها المقوِّم النوعي هي وصفية Descriptive وتفصيلية عن الموقف الملاحظ.

إن استخدام الملاحظة في الدراسات التقويمية متنوعة وتتراوح بين الملاحظة بالمشاركة Participant Observation والملاحظة المقيدة Structured Observation فالملاحظة بالمشاركة ترجع جذورها الى الانثروبولوجيين القدماء الذين استخدموها في دراسات الثقافات البشرية من أمثال (مارجريت ميد). أما الملاحظات المفيدة فتستخدم بشكل واسع في تقويم البرامج والنشاطات الطلابية التي بصعب تقويمها بشكل خاص في الاختبارات التحريرية وانواع اخرى من بحوث التقويم الكمية.

كما يمكن استخدام الملاحظة بجمع بيانات تدعم أو تكمل بيانات جمعت بطرق اخرى، كالمقابلة او من خلال تحليل الوثائق، فاذا كانت المقابلة اسلوباً هاماً في دراسة تقويمية معينة، فان الملاحظة توفر أدلة تؤكد ما تم التوصل اليه عن طريق المقابلة. واذا حصل مقوم ما على بيانات جمعت عن طريق استبيان وزع على مجموعة من المستفيدين أو العاملين في برنامج معين وشك في مصداقيتها، فان الملاحظة يمكن أن تدعم أو تنقض تلك البيانات لانها تجمع

من الواقع المدروس بشكل مباشر وطبيعي وواقعي. وفي كثير من الأحيان تكون الملاحظة هي الوسيلة الوحيدة في جمع البيانات كما في دراسات الظاهراتيه والأثنوغرافيا Phenomenology and Ethnography.

تعريف الملاحظة:

تعني الملاحظة بمعناها البسيط الانتباه العفوي إلى حدث أو ظاهرة أو أمر ما، وتعرف أيضاً بأنها: الاعتبار المتنبه لحادثةٍ أو ظاهرةٍ أو شيء ما.

أما التعريف العلمي للملاحظة فهي أنها انتباه مقصود ومنظم ومضبوط للظواهر أو الأحداث أو الأمور بغية اكتشاف أسبابها وقوانينها. أو هي عملية مراقبة أو مشاهدة السلوك والظواهر والمشكلات والأحداث ومكوناتها المادية والبيئية ومتابعة سيرها واتجاهاتها والعلاقات بين عناصرها بأسلوب علمي منظم ومخطط وهادف بقصد التفسير وتحديد العلاقة بين المتغيرات والتنبؤ بتطور الظاهرة وتوجيهها لخدمة الإنسان وتلبية احتياجاته.

أنواع الملاحظة:

تصنف الملاحظة تصنيفات عدة وفق ظروف خاصة وضوابط معينة ومن تلك التصنيفات ما يأتي:

أولاً: تصنيف الملاحظة حسب الضبط: وفيه تصنف الملاحظة إلى نوعين:

أ. ملاحظة بسيطة: **Simple Observation**

وتحدث عندما يُلاحظ الباحث الظواهر والأحداث كما تحدث طبيعياً وفي ظروفها العادية دون إخضاع تلك الظاهرة للضبط العلمي، أي دون إعداد مسبق ودون إستخدام أدوات التسجيل أو التصوير، ويفيد هذا النوع في الدراسات الاستطلاعية الأولية التي يحاول فيها الباحث جمع بيانات أولية عن الظواهر والأحداث تمهيداً لدراستها دراسة دقيقة في المستقبل.

ب. ملاحظة منظمة: **Systematic Observation**

وهي التي تخضع لدرجة عالية من الضبط العلمي، حيث يحدد الباحث

الأدوات الممكن استخدامها أثناء الملاحظة والمكان الذي ستقع فيه الملاحظة، وتحديد العينات الزمنية التي ستجري خلالها الملاحظة، وكذلك تحديد أسلوب تدوين المعلومات أي هل يتم التدوين خلال الملاحظة أم بعد الانتهاء منها، كما يتم تحديد الأشخاص الذين ستجري ملاحظتهم وفقاً لطبيعة الموضوع وأهدافه بما يخدم توفير إجابات دقيقة لأسئلة القائم بالملاحظة.

ثانياً: من حيث عدد القائمين على الملاحظة:

وتصنف إلى نوعين هما :

أ. الملاحظة الفردية:

وهي التي يقوم بها شخص واحد لجمع بيانات معينة عن الموقف الملاحظ.

ب. الملاحظة الجماعية :

وهي التي يقوم بها أكثر من شخص واحد لملاحظة موقف معين ، ويمكن أن تكون أكثر دقة وموضوعية لا سيما عندما تتطابق أو تتقارب نتائجها بين الملاحظين، كذلك إذا نسى الملاحظ معلومة معينة فإن الملاحظ الآخر أو الملاحظين الآخرين المشاركين معه يتذكرونها، وبذلك نضمن تسجيل معظم جوانب الموقف الملاحظ.

ثالثاً: تصنيف الملاحظة حسب دور القائم بالملاحظة:

حيث تصنف إلى نوعين هما :-

أ. الملاحظة بالمشاركة: Participant Observation

يعرف كل من بوجدان وتايلر (Bogdan & Taylor, 1975) هذا النوع بأنه فترة مكثفة للتفاعل الاجتماعي بين الباحث ومجموعة أفراد في بيئة معينة يتم من خلالها جمع معلومات منظمة وهادفه، أما جوبا ولنكولن (Guba & Lin-coln,1981) فيعرفانها بأنها شكل من أشكال الاستقصاء يلعب فيها الباحث نوعين من الأدوار، في الأول يقوم بالملاحظة، أي أنه يقوم بملاحظة الموقف

من الخارج، وفي الثاني يصبح فيه عضواً مشاركاً في المجموعة ويكون جزءاً ورمزاً من نشاطاتها ونتائجها، ويضيف كروس ويل (Creswell,2002) دوراً ثالثاً للملاحظ ذلك هو غير المشارك والمشارك في نفس الوقت، حيث بعض الحالات تتطلب من المقوّم أن يلاحظ الموقف عن بعد أولاً أي بصيغه غير مشارك وذلك لجمع بعض المعلومات التي توضح له صورة الموقف وظروفه ثم يتبعها بدور المشارك في الجماعة التي تجري ملاحظتها، فلو أراد باحث أن يجمع معلومات تفيده في دراسة عن التفاعل الاجتماعي، أو نوع الأحاديث التي يتبادلها طلاب الجامعة خلال فترات الاستراحة، فبعد أن يحدد المكان والزمان الذي ستجري فيه الملاحظة وهو النادي الطلابي خلال فترة الغداء، يحاول أولاً أن يراقب المكان في الزمان الذي حدده لمرات متعددة ليلاحظ سلوك الطلاب خلال تلك الفترة والوقت الذي يستغرقة الطلبة في النادي، واسلوب دخولهم وخروجهم وكيفية انتشارهم بشكل جماعات أم فرادى، إلى أن تتوفر لديه معلومات كافية عن تلك الأمور، وهو هنا ملاحظ عن بعد(غير مشارك). ثم يحول دوره إلى ملاحظ بالمشاركة. وذلك عندما يجلس مع مجموعة من الطلاب على مائدة واحدة ويتناول الغداء بينهم، ثم يستمع إلى أحاديثهم ويشاركهم، ويديرها أحياناً. وعند ذاك يصبح جزءاً من الموقف وعضواً فيه ليسمع ويرى كل السلوكات التي يمارسها الطلاب في فترة الغداء في النادي الطلابي ويجمع المعلومات الخاصة بالأحاديث التي يتبادلوها أو السلوكات التي يمارسوها وبشكل مباشر وعند ذاك يكون الملاحظ قد مارس أولاً الملاحظة بدون مشاركة ثم انتقل إلى دور الملاحظ بالمشاركة.

فوائد الملاحظة بالمشاركة:

هناك العديد من الفوائد التي يجنيها المقوّم والباحث من هذا الأسلوب ومنها:

1- في هذا الأسلوب يتم تطوير خبرات مباشرة عن الموقف الملاحظ، ويشير دوجلس (Douglas,1976) أننا في حياتنا اليومية نستخدم اختيارات متنوعة

للتأكد من الحقائق، ولكن معظم تلك الاختيارات المهمة تأتي من خلال الخبرات المباشرة التي تشكل فيها الملاحظات عنصراً أساسياً ، وأن الخبرة هي أفضل معلم، وأن الأفراد الذين يستفيدون من تراكم الخبرة يكون بإمكانهم استخدامها في أي وقت كما أن الملاحظة بالمشاركة هي بلا شك أفضل أسلوب يؤدي إلى معرفة الحقائق كما هي، بل وكما يدركها الشخص الملاحظ عن قرب، لذلك تكون هذه الحقائق أكثر دقةً وأكثر موضوعية.

2- تفيد الملاحظة بالمشاركة في بناء أدوات القياس التي نريد تطويرها، فمن خلال مشاهداتنا لطفل يسلك سلوكاً معيناً ـ على سبيل المثال- ستنشأ لدينا خبرات وخلفيات كثيرة يمكن الاستفادة منها في بناء فقرات أو أسئلة أو تقديرات لتحديد صفة أو خاصية سلوكية معينة نريد دراستها من خلال أداة قياس معينة .

3- إن الملاحظة بالمشاركة تمدنا بالكثير من المعلومات التي تساعدنا على فهم بعض المواقف المتعلقة بالمعتقدات والاتجاهات والعادات الاجتماعية والدوافع الإنسانية ، وجدير بنا هنا أن نتذكر أن الكثير من علماء الأنثروبولوجيا توصلوا إلى نظريات وتفسيرات للكثير من السلوكيات الاجتماعية في العديد من المجتمعات الإنسانية نذكر فيها على سبيل المثال الدراسات الرائدة للعالمة(ميرغريت مي) في المجتمعات البدائية.

4- في الملاحظة بالمشاركة يكون بإمكان الباحث جمع بيانات غزيرة عن الموقف الملاحظ دون الانتظار الطويل الذي تتطلبه عملية توزيع وجمع الاستبيان ـ مثلاً ـ وما تستغرقه عملية تفريغه وتحليله وتفسير نتائجه من وقت وجهد كبيرين.

ب ـ الملاحظة بدون مشاركة :

وفيها يلعب الباحث دور المشاهد أو المتفرج بالنسبة للظاهرة أو الحدث موضوع الدراسة بحيث يكون بعيداً عن مسرح الموضوع الواقع تحت الملاحظة قدر الإمكان، ويمتاز هذا النوع بالموضوعية على أساس أن ابتعاد الباحث عن

الظاهرة يقلل من تأثره بها، لكن ما يؤخذ على هذا الأسلوب أن الباحث الملاحظ لا يستطيع أن يلم بجميع جوانب الموضوع الملاحظ وهو بعيد عنهُ.

أهمية الملاحظة في التقويم:

يمكن تلخيص دور وأهمية الملاحظة في عملية التقويم التربوي في الجوانب الآتية:

1- توفر الملاحظة المباشرة للمقوِّم فرصة مناسبة للإطلاع على الكثير من الامور التي ربما لا ينتبه لها العاملون في البرنامج على اساس انها مألوفة لديهم ويلاحظونها كل يوم.

2- تمكن الملاحظة المقوِّم من فهم السياق العام الذي يُنفذ البرنامج في إطارهِ وهذا أمر مهم وضروري لتكوين المنظور الشامل للبرنامج.

3- يستطيع المقوِّم من خلال الملاحظة المباشرة للبرنامج التأكد من دقة وصحة التقارير التي يصدرها مسؤولو ذلك البرنامج بعد اجراء مقارنة بين تلك التقارير وبين واقع الحال الذي يلاحظه المقوم بشكل مباشر.

4- يستطيع المقوِّم عن طريق الملاحظة المنظمة التعرف على معلومات كثيرة عن البرنامج ربما لا يمكن توفيرها عن طريق المقابلة لا سيما المتعلق منها بأمور حساسة كالقضايا المالية أو اسلوب المسؤول في التعامل مع العاملين في البرنامج المقوَّم لحساسية الحديث في مثل هذه الامور اثناء المقابلة.

5- تسمح الملاحظة للمقوِّم الفرصة لاستخدام معارفه وخبراته الشخصية لفهم وتفسير عمل البرنامج المقوَّم، لا سيما وان تلك الخبرات والانطباعات تصبح وسائل مساعدة هامة تستخدم لفهم طبيعة عمل البرنامج.

6- تساعد الملاحظة المباشرة المقوِّم على فهم العلاقات الانسانية بشكل أوضح سواء بين المسؤول والعاملين أو بين العاملين أنفسهم ضمن البرنامج الواحد.

7- توفر الملاحظة الفرصة للقائم بالملاحظة من رصد حسن أو سوء استخدام الموارد المالية في تطوير اساليب العمل وتوسيعه وتوفير التقنيات اللازمة لتسريع العمل او زيادة وسائل الانتاج أو كيفية استخدام جزء من الموارد

كمحفزات مادية للعاملين مؤثرة في رفع مستوى رضائهم عن العمل وادائهم له.

عيوب الملاحظة:

1- إن بعض المفحوصين حين يشعرون بأنهم ملاحظون قد يغيرون من سلوكهم ولا يظهرون السلوك المستهدف، ويمكن تقليل أثر هذا العيب بتكرار الملاحظة وبقاء الباحث مع المفحوصين مدة أطول حتى يألفوه ولا يبقى لوجوده أي أثر في تعديل سلوكهم .

2- تتطلب الملاحظة وقتا طويلا، فقد ينتظر الباحث فترة طويلة حتى يبرز السلوك الذي يلاحظه، فإذا أراد الباحث ملاحظة سلوك الطالب حين يعاقبه المعلم، فإنه يدخل إلى غرفة الصف وينتظر حتى يعاقب المعلم الطالب، وقد لا يحدث ذلك السلوك مما يؤدي إلى تكرار الملاحظة وربما لا ترى السلوك المطلوب.

3- أسلوب الملاحظة يخضع عادة لفترة زمنية معينة، فالأحداث التاريخية، لا يمكن إخضاعها للملاحظة.

4- بعض الأحداث لا يمكن استخدام الملاحظة فيها كالحياة الخاصة للأفراد، ولكن يمكن الحصول على المعلومات بشأنها بواسطة المراسلة أو المقابلة الشخصية، أو الاستبيانات.

5- قد تعيق عوامل منظورة عملية القيام بالملاحظة كتقلبات الطقس مثلا، أو وقوع أحداث أخرى بديلة.

6- تعذر تسجيل الملاحظة من خلال الملاحظة بالمشاركة الكاملة، مما يجعل الملاحظ يعتمد على ذاكرته في وصف الوقائع، ولهذا الأمر آثاره السلبية على القيمة العلمية للمعلومات ودقتها.

7- احتمال تغير السلوك إلى ما يريده الملاحظ نتيجة لمشاركته الكاملة، إذ قد يوجه الأفراد توجيها لا إراديا إلى ما يبحث عنه، مما يتنافى مع ما تتطلبه الملاحظة العلمية من تجرد وعدم تأثير، وهذا يظهر في الملاحظة بالمشاركة.

قواعد استخدام الملاحظة

لكي تكون الملاحظة دقيقة وصادقة يجب اتباع القواعد الآتية:

1- يجب التخطيط مسبقا لما يجب ملاحظته وذلك بناء على أهداف محددة نريد قياسها، مع الفصل تماما بين ما نلاحظه وأية تفسيرات أخرى يمكن استنتاجها من الملاحظة.

2- يجب أن يركز الباحث على نوع أو نوعين من السلوك فقط وعدم تشتيت ذهنه لاكثر من ذلك.

3- يجب استخدام صفات واضحة غير غامضة حتى تكون الملاحظة محددة تصف السلوك وصفا سليما.

4- يجب أن يكون كل سلوك ملاحظ مختلف تماما عما عداه من أنواع السلوك الأخرى، ذلك لأن تداخل أنواع من السلوك يجعل من الصعب وصفها وتصنيفها.

5- يجب أن يكون الباحث واعيا لأخطاء الملاحظة التي قد تحدث نتيجة لإختيار أوقات غير مناسبة لملاحظة ذلك السلوك.

6- يجب تسجيل وتلخيص الملاحظات عقب حدوثها مباشرة او تسجيلها اثناء الملاحظة في الاستمارة المعدة لهذا الغرض.

7- يجب تأجيل تفسير السلوك إلى ما بعد الانتهاء من الملاحظة فكثيرا ما يترتب على محاولة التفسير أثناء حدوثه الإخلال بموضوعية الملاحظة.

8- يجب ألا يظهر الباحث أنه يلاحظ سلوكا ما أو فردا ما حتى لا يشعر الملاحظ أنه تحت المراقبة.

9- لا بأس أن يتأكد المقوم من صدق ملاحظاته من خلال مقارنتها مع ملاحظ آخر أو القيام بالملاحظة أكثر من مرة اذا كان ذلك ممكناً.

10- يمكن الاستفادة من اجهزة التسجيل المرئي أو الصوتي لتسجيل الموقف الذي تتم ملاحظته.

شكل (5-1)

مخطط بسيط لاستمارة ملاحظة موقف معين

● مكان الملاحظة

- الشخص الملاحظ أو (الحدث الملاحظ)
- الشخص القائم بالملاحظة:
- دور القائم بالملاحظة:

(مشارك في الموقف الملاحظ؟..................ملاحظ عن بعد (غير مشارك)؟.......)

×الوقت الذي تمت فيه الملاحظة؟ من الساعة..............حتى الساعة .... اليوم.................. التاريخ.................

● وصف الملاحظات أي وصف الاحداث في الموقف الملاحظ حسب التسلسل الزمني لحدوثها، لا بأس من رسم صور تخطيطية للأفراد الذين كانوا. الوضع المادي للموقف الملاحظ، الاحداث والنشاطات.

● ملاحظات خاصة بالقائم بالملاحظة.

(ملاحظات عن خبرات الملاحظ، الحس الداخلي والتصورات وافكاره حول الموضوع الملاحظ)

● المصدر: (Creswell, 2001)

المقابلة: Interview

هي: استبيان شفوي يقوم من خلاله الباحث بجمع معلومات وبيانات شفوية من المفحوص (الشخص الذي يقابله) مباشرةً. ويمكن تعريفها أيضاً بأنها محادثة تتم وجهاً لوجه بين القائم بالمقابلة Interviewer والمستجيب Re-spondent، بغرض الحصول على معلومات محددة عن موضوع مقصود.

وتستخدم المقابلة كأسلوب لجمع البيانات والمعلومات في صور شتى ولتحقيق اهداف متنوعة فالمعلم مثلاً يجري العديد من المقابلات خلال يومه المدرسي كمقابلة العديد من طلابه خارج حجرة الصف في مكتبه على وجه التحديد بغية التعرف على مشكلاتهم أو تقديم الارشادات والنصائح او بغرض

مساعدتهم من خلال اجابة استفساراتهم نحو المادة الدراسية وحتى مشكلات فردية قد يعانون منها في المدرسة او خارجها ومن بعض المشكلات العائلية.

كما أنه يجري في كثير من الأحيان مقابلات مع أولياء أمور الطلاب سواء لاطلاعهم على نقاط القوة أو الضعف في المسيرة الدراسية لابنائهم، او لايصال بعض معاناة الابناء من الجو العائلي او المساهمة في حل المشاكل العائلية التي تواجه الطلبة من خلال الاستفسار من الآباء عن جذور وظروف تلك المشكلات للتوصل الى حلول تخفف منها أو قد تزيلها بشكل نهائي، ومن هنا يتطلب هذا الأمر من المعلم نوعاً من الكياسة واللباقة والقدرة على التأثير بالآخرين أي معرفة المعلم لمستلزمات المقابلة واصول اجرائها.

كما أن المقابلات يمكن أن توفر الكثير من المعلومات الضرورية لتقويم البرامج والمشروعات والمؤسسات التربوية عندما يجريها المقوِّم مع المسؤولين عن تلك البرامج والمشروعات والمؤسسات التربوية والمستفيدين منها.

وتتميز المقابلة بالمرونة واتساع نطاق استخداماتها، فالقائم بالمقابلة يمكنه أن يغير نمط تساؤلاته إذا تطلب الأمر ذلك ويعيد صياغة السؤال إذا لم يكن واضحاً للشخص المقابَل، كما تفيد المقابلة مع الاطفال والاشخاص الذين لا يعرفون القراءة والكتابة. وتتطلب المقابلات بمختلف أنواعها إقامة علاقة إنسجامية مع المستجيب من أجل كسر حالة الشك والتردد وتحفيزه لكي يكون صريحاً ومنفتحاً مع الشخص الذي يجري المقابلة فيزوده بالمعلومات التي يرغب بها، كما أن الظروف المكانية والزمانية لها دور كبير في هذا الامر، فلا بأس أن تتم المقابلة في مكان عمل الشخص الذي نريد مقابلته لأن ذلك يعطيه نوعاً من الشعور بالأمن والطلاقة اكثر من أن نطلبه للمقابلة في مكان آخر يتصف بالرسمية والشكلية بما يزيد من تردد المستجيب ويقلل من تفاعله مع المقوِّم. وعلى الرغم من أن المقابلة تستخدم في العديد من مجالات البحث والتقويم لكنها قد تكون الوسيلة الاساسية في أنواع محددة ومنها البحث والتقويم النوعي وخصوصاً في الاثنوغرافيا والظاهراتية.

خطوات إعداد المقابلة:

قبل البدء بإجراء المقابلة لا بد للقائم في المقابلة التخطيط لها بشكل دقيق لا سيما في الجوانب الآتية:

1- **تحديد أهداف المقابلة:** على الباحث تحديد الأهداف التي من أجلها يتم إجراء المقابلة وطبيعة المعلومات التي يحتاجها لمشكلة دراسته وصياغتها بالشكل المناسب.
2- **تحديد الأفراد الذين سيقابلهم:** أي تحديد المجتمع الأصلي للدراسة واختيار عينة ممثلة تحقق أغراض الدراسة بشرط توفر الرغبة عند أفراد هذه العينة لإعطاء المعلومات المطلوبة.
3- **تحديد أسئلة المقابلة:** ضرورة الإعداد المسبق للأسئلة حتى تتوفر فيها مزايا الأسئلة الواضحة والمقبولة من قبل المفحوص.
4- **تحديد مكان وزمان المقابلة:** ويراعى في ذلك المكان المريح والمقبول وكذلك أن لا يتعارض الوقت مع وقت المفحوص.
5- **تحديد الطريقة التي يستخدمها المقوِّم** لتسجيل أجوبة المستجيبين وهل تكتب مباشرة أم تؤجل الى ما بعد الانتهاء من المقابلة، وهل بالامكان استخدام الاجهزة الصوتية والمرئية لتسجيلها.

تنفيذ المقابلة:

- ضرورة التدريب على إجراء المقابلة: أي قيام الباحث باختيار عينة صغيرة من زملائه ليجري معهم مقابلات تجريبية ويحدد النقاط الايجابية والسلبية في إجراءاته.
- التنفيذ الفعلي للمقابلة: ينفذ الباحث المقابلة مع العينة المطلوبة مراعيا الآتي:

- البدء بحديث مشوق ومتدرج لتوضيح أهداف المقابلة وتوضيح الدور المطلوب من المفحوص.
- إظهار الود للمقابل بحيث يشعر بالأمان والهدوء الذي يدفعه لإعطاء الاستجابات المطلوبة.

- البدء بمناقشة الموضوعات المحايدة التي لا تحمل صبغة انفعالية أو شخصية حادة لدى المفحوص.
- طرح الأسئلة بوضوح ولا مانع من التوضيح عند الضرورة.
- إعطاء الوقت الكافي للمفحوص لتقديم الإجابة المطلوبة.
- توجيه المفحوص نحو الالتزام بالسؤال وحصر الحديث بالاتجاه الذي يريده الباحث وأعادته بلباقة عندما يخرج عن موضوع المقابلة.
- عدم إظهار الباحث أي انفعال أو دهشة لسماعه معلومات محددة أو استنكاره لموقف معين حتى لا يؤثر على استجابات المفحوص.
- عدم إحراج المفحوص واتهامه أو توجيه أسئلة هجومية عليه.

أنواع المقابلات:

هناك تصنيفات عدة وضعها الباحثون في مجال أدوات التقويم للمقابلة منها:

(1) تصنيف المقابلة وفقاً لبناؤها أو درجة تقنينها:

أ - المقابلة المحددة البناء أو المقابلة المقيدة Structured :

وتكون اجراءاتها ومحتواها مقننة مسبقاً أي يتم بناء خطواتها وفقاً لمعايير احصائية وتطبيقية بحيث يتم الوصول بعد ذلك الى صياغة الشكل النهائي لما تتضمنه من فقرات أو اسئلة ذات اجابات محددة باختيارات للاجابة لا يخرج عنها المقابِل والمستجيب، مثال ذلك: هل تعتقد أن أهداف البرنامج الحالي محددة بشكل واضح ودقيق؟ فيكون جواب المستجيب اختيار احد الاستجابات التالية: (نعم، لا، محددة الى حدٍ ما).

ويستخدم هذا النوع من المقابلات في التقويمات الكمية التي يجري تحليل بياناتها وفقاً لمعاملات احصائية وصفية أو استدلالية.

ب - المقابلات غير المقيدة Unstructured :

وهي مقابلات تشبه الى حدٍ ما الاستبيان المفتوح لكنها تختلف عنه في أن

اجرائها يكون بشكل شفوي لا كتابي حيث تتطلب اجابات حرة من المستجيب وتساعد القائم بالمقابلة بانشاء اسئلة اخرى أو تعديل في اسئلة سابقة، لكن ما يعاب عليها اختلاف الباحثين في تفسير الاجابات من جهة وصعوبة تكميم تلك الاجابات من جهة أخرى. مثال: (هل تعتقد أن أهداف البرنامج الذي تعمل في ظله محددة بشكل واضح ودقيق؟).

(2) تصنيف المقابلة حسب طبيعة الهدف المرجو منها:

أ- مقابلات تقويمية:

وهي التي تهدف الى التوصل لمدى نجاح أو فشل برامج أو مشروعات تربوية أو تحسينات جديدة أُدخلت على برامج معينة لتطوير ادائها وذلك من خلال التعرف من المستجيبين على نقاط الضعف والقوة وآليات التطوير المقترحة.

ب- مقابلات فردية للمسؤولين في موقع اتخاذ القرارات:

حيث يمكن أن يجري المقوِّم هذا النوع مع القيادات الإدارية العليا في دوائر رسمية أو مؤسسات غير رسمية للحصول على استجابات تتعلق بطبيعة عمل برنامج معين من حيث أهدافه وأساليب تحقيقها وآلية العمل ونوع المخرجات وبعض نقاط القوة والضعف وأسبابها وأساليب معالجتها، ومدى انسجام القرارات المهمة المتخذة بشأن البرنامج مع طبيعة عمل البرنامج من أجل تقويم فاعليتها وأهميتها في زيادة كفاءة ذلك البرنامج، وكثيراً ما تكون هذه المقابلات من النوع المفتوح.

(3) تصنيف المقابلة وفقاً لعدد الأشخاص الذين تتم مقابلتهم:

أ- مقابلات فردية:

حيث يقابل بها المقوِّم فرداً واحداً على انفراد يعقبه فرد آخر وهكذا، ويُجرى هذا النوع عندما يريد المقوِّم ان يكون المستجيب لوحده ليتمكن من إبداء رأيه حول الموضوع المقوَّم بحرية كاملة دون تأثير الأفراد الآخرين العاملين معه في العمل.

ب- مقابلات جماعية:

وفيها تتم مقابلة مجموعة قد تكون كبيرة من الافراد العاملين في مكان واحد مرةً واحدة للتعرف على وجهة نظرهم المشتركة حول الموضوع الذي يتم تقويمه، حيث نطرح العديد من الافكار والملاحظات التي تساعد المقوِّم على استنتاجات هامة لجوانب متعددة من البرنامج.

أسئلة المقابلة **Questions of interview**

يَعتمد نجاح المقابلة والحصول على المعلومات من المستجيب لأغراض التقويم على عدة أمور منها ما يتعلق بشخصية القائم بالمقابلة والظروف الزمانية والمكانية التي تجري في ظلها المقابلة وامور اخرى تتعلق بالمستجيب نفسه من حيث تعاونه ومدى تقديمه للمعلومات، لكن أحد ابرز العناصر الهامة في تحقيق المقابلة لاهدافها هو أسلوب طرح السؤال على المستجيب، ونوع السؤال ومدى قدرته على استخراج المعلومة المطلوبة من المستجيب لا سيما عندما يطرح في مكانه ولا يُثير في المستجيب أي انفعال أو قلق أو خوف، اذ من المعلوم أن الاسئلة المستخدمة في عملية التقويم حساسة لأنها تتطلب في غالب الاحيان التأشير على جوانب الخلل في جوانب عدة من البرنامج، بل وتمتد أحياناً الى نقد المسؤولين عن ادارة البرنامج، مما يُثير لدى بعض المستجيبين الخوف والتردد. ولكي نلم بأهمية استخدام السؤال في المقابلة بالشكل الصحيح لا بد من التعرف على أنواع أسئلة المقابلة ومتى وأين يستخدم كل نوع، وكما يأتي:

1- الأسئلة التمهيدية:

وتهدف هذه الأسئلة الى تشجيع المستجيب على إعطاء أكبر قدرٍ من المعلومات المتعلقة بالموضوع، حيث يتابع القائم بالمقابلة علاقة المستجيب بالموضوع ودوره من حيث أنه عامل في البرنامج المقوَّم أم مستفيد وما تجاربه الشخصية وملاحظاته عنه والاستماع إليه بإهتمام، كأن نسأله السؤال الآتي:

باعتبارك أحد العاملين في هذا البرنامج وبالتأكيد لديك ملاحظات ووجهات نظر شخصية عن البرنامج فهل لك أن توضح لنا تلك الملاحظات ووجهات النظر؟

2- اسئلة المتابعة:

وتطرح هذه الاسئلة بأشكال مختلفة يفهم منها المستجيب أننا نريد منه أن يستمر في حديثه، وتتضمن كلمات مقتضبه أو اشارات مثل: جيد، نعم، أو هز الرأس أو أية اشارة جسدية كعلامة للانتباه والمتابعة، ويتطلب استخدام ذلك إنتباه القائم بالمقابلة لما يقوله المستجيب وزيادة هذه المتابعات عندما يطرح أموراً هامة توفر معلومات جيدة للموضوع قيد التقويم.

3- أسئلة التمحيص:

وتهدف للحصول على اجابات متعمقة يريد المقابل التدقيق في محتواها ومن الأمثلة على ذلك: هل بالامكان أن نتوقف عند الموضوع أو النقطة التي ذكرتها، هل لديك تفصيلات أكثر عنها؟ هل من مثال على الجانب الذي ذكرته؟ هل كان هناك آخرين كانوا معك في الموقف الذي ذكرته؟

4- الأسئلة غير المباشرة:

وتهدف الى استكشاف بعض المواقف الخاصة عن برنامج معين، حيث لا يتم طرحها بشكل يتعلق بموقف الفرد مباشرةً، مثال ذلك عندما نسأل طالباً معيناً عن رأيه في المنافسة على العلامات من قبل الطلبة؟ فبدلاً من أن نسأله عن رأيهُ أو موقفه حول الموضوع، يمكن توجيه السؤال بالشكل الآتي:

- ما هو رأي أولياء الامور بتعيين مجموعة من المعلمين الذين يزاولون المهنة لأول مرة في هذه المدرسة؟

5- الاسئلة المباشرة:

وتتمحور هذه الاسئلة حول أبعاد ونواحٍ محددة عن الموضوع المراد تقويمه، ولذلك يُفترض تأخير هذا النوع لمرحلة لاحقة وليس في بداية المقابلة وذلك بعد أن يكيف المقابل الظروف النفسية المحيطة بالمستجيب وخاصةً بعد أن تكلم عن امور أراد أن يتكلم عنها حيث تصبح عند ذاك الفرصة سانحة

للمقوِّم بطرح مثل هذه الاسئلة الجوهرية المباشرة حول الموضوع. مثال ذلك:

أ- هل بالامكان تحديد جوانب الخلل التي أدت الى صعوبات حالت دون تحقيق البرنامج لأهدافه رغم التخصيصات المالية الكافية؟ مثال آخر يتعلق بالتقويم الذاتي لمدرسة معينة:

- هل تعتقد أن المعلمين متحمسين لاجراء التقويم الذاتي في هذه المدرسة؟
- هل لدى المعلمين وادارة المدرسة الخبرة الكافية لاجراء عملية التقويم الذاتي؟

6- الأسئلة التفسيرية:

وتتطلب إعادة صياغة اجابة عن سؤال معين. مثل: ماذا تقصد بإهمال المعلمين لأعمال الطلاب؟ أو: هل تقصد أن المعلمين لا يعطون اهتماماً كافياً لاعمال الطلاب؟. أن هذه الاسئلة تهدف التأكد مما قاله المستجيب، أو التأكد من تفسير المقوِّم واستنتاجه باقوال المستجيب. وأخيراً لا بد من الاشارة الى ضرورة استخدام القائم بالمقابلة "للصمت" أثناء المقابلة حتى يستطيع اعداد صياغة جيدة لاسئلة تقفز الى ذهنه ويراها ضرورية، كما انه يوحي للمستجيب أن المقوِّم يهتم ويصغي لأقواله، لكن ذلك يجب أن لا يستغرق فترات طويلة تُضيع فرصاً كثيرة من تدخل الباحث لتوجيه المقابلة الى الموضوع الرئيسي. (أبو زينة وآخرون، 2005).

عيوب المقابلة:

1 - قد يتعرض تفسير نتائجها للعوامل الذاتية للمقوِّم وتمتزج بتصوراته الشخصية.

2- تستغرق وقتا طويلا في تنفيذها.

3- ذات تكلفة مادية وزمنية مرتفعة.

4- تحفظ بعض المستجيبين عن تقديم المعلومات المطلوبة التي يسعى المقوِّم للحصول عليها.

5- اعطاء بعض المستجيبين اجابات مرغوبة لدى القائم بالمقابلة أو الاجابة المرغوبة اجتماعياً social desirability لا سيما في الموضوعات المتعلقة بالمعايير أو التقاليد الاجتماعية وبذلك تضيع فرصة الحصول على الاستجابة الشخصية المتفردة التي يود المقوِّم الحصول عليها.

ملاحظات عامة

دلت الدراسات العملية المتعلقة بالمقابلة على ملاحظات عديدة منها ما يأتي:

1- يفضل ان يكون عمر الباحث مساويا أو اكبر من عمر المستجيب.
2- أن المتخصص في العلوم السلوكية أفضل في اجراء المقابلة من المتخصص في العلوم الطبيعية.
3- كلما زادت خبرة الباحث وتدريبه زادت دقة الاجراءات ونجاح المقابلة.
4- يمكن القول إن المؤهل العلمي للقائم بالمقابلة مهم في نجاحها، فمثلاً خريج الجامعة افضل من غير الخريج في ادارة المقابلة وخاصةً بعد تدريبه على اصول اجرائها.
5- إن الشخص الذي لديه معرفة وتجارب سابقة بأساليب إقامة العلاقات الاجتماعية أكثر قدرة من غيره في ادارة الحوار مع المستجيب والحصول على المعلومات المطلوبة منه.

شكل (2 - 5) مخطط المقابلة

- وقت المقابلة:

الزمن المستغرق: من الساعة.......................الى الساعة...............

اليوم.......................التاريخ..................

المكان...

- اسم القائم بالمقابلة...................................
- اسم الشخص الذي تمت مقابلته..............................
- وظيفة القائم بالمقابلة..................................
- وصف المقابلة:

(وصف الموضوع، اخبار المستجيب عن: هدف المقابلة. الافراد والمصادر التي سيتم جمع البيانات منهم. ماذا سيتم عمله بالبيانات. واشعار المقابل بان المعلومات ستبقى سرية كاملة، الوقت الذي ستستغرقه المقابلة، قراءة اجابات الشخص وتوقيعه عليها. (يمكن اعادة تسجيل الاجابات بجهاز التسجيل ان كان مستخدما).

أمثلة لاسئلة مقابلة:

1- الرجاء وصف دورك في الحادثة التي وقعت بتاريخ...؟

2- ماذا حدث لك منذ الحادثة التي شاركت فيها؟

3- ماذا تتوقع ان يقع على المجتمع من اذى نتيجةً لهذه الحادثة؟

4- ما العواقب التي يمكن ان تنتج عن هذه الحادثة؟

5- من هم الأشخاص الذين تقترح أن نتحدث معهم حول هذا الموضوع ليزودونا بمعلومات مفيدة؟

ملاحظة: إشكر الشخص الذي تمت مقابلته، واكد له ان المعلومات التي افاد بها ستبقى موضع ثقة وسرية كاملة، واحتمال اجراء مقابلات اخرى معه في المستقبل.

المصدر: (Creswell , 2002)

قوائم المراجعة (الشطب) Checklists

كثيراً ما يستخدم المعلمون والمرشدون والموجهون التربويون هذه القوائم لرصد وتقويم العديد من جوانب العملية التربوية سواء المتعلقة منها بالطالب أو المعلم أو الكتاب المدرسي أو المهارات التي يظهرها الطالب في نشاطات أخرى، لا سيما تلك التي لا يتلائم معها استخدام الاختبارات التحصيلية. فعند وصف سلوك الطالب في الصف أو المختبر أو في نشاط ما خارج الصف

تصلح هذه القوائم في تحديد ما يظهره أو لا يظهره الطالب من جوانب السلوك المطلوب تقويمه.

وفي تقويم البرامج التربوية يستطيع المقوِّم استخدامها للحكم على جانب أو جوانب محددة من عمل ذلك البرنامج لأنها تعطي صورة أكثر وضوحا عن البرنامج من خلال تحديد ما هو موجود وما هو غير موجود من مظاهر أو مؤشرات هامة في إصدار الحكم عليه، كما يمكن استخدام هذه القوائم في عملية التقويم الذاتي لفرد أو برنامج أو مؤسسة تربوية معينة.

وتتألف هذه القوائم من مجموعة من الأسئلة أو الفقرات تكون الإجابة عنها بكلمات محددة مثل (نعم، لا، موجود، غير موجود . . . إلخ) يؤشر المقوِّم على إحداهما كدليل على السلب أو الإيجاب، وقد يكتفي بوضع إشارة (√) تحت استجابة محددة، وفيما يلي نموذج لقائمة مراجعة خاصة بتقويم التفاعل الاجتماعي للطالب مع زملائه الآخرين.

اسم الطالب: ________________

اسم المقوِّم ________ تاريخ التقويم ________

التعليمات:

ضع إشارة (√) تحت الاستجابة التي تمثل ما تلاحظهُ أو ما لم تلاحظه من أنماط السلوك الاجتماعي المتمثل بكل فقرة من الفقرات المدرجة في أدناه:

م	الفقـــــرات	نعم	لا
1	يُشارك الآخرين من جنسه استخدام أدوات الدراسة.		
2	يشارك الآخرين من الجنس الآخر استخدام أدوات الدراسة.		
3	نادرا ما يشارك أي طالب آخر في استخدام أدوات الدراسة.		
4	يظهر صعوبة في تفاعله مع الآخرين.		
5	يساعد الآخرين في حل مشاكلهم.		
6	يقاطع الآخرين أثناء حديثهم.		
7	يفسح المجال للآخرين أثناء اللعب أو العمل.		
8	يستمع باهتمام لآراء الآخرين.		
9	يتعصب لآرائه ويعارض آراء الآخرين.		
10	يحاول إيذاء الآخرين أثناء اللعب أو العمل.		
11	لديه نزعة السيطرة على الآخرين.		
12	يحترم بشكل واضح زملائه الآخرين.		
13	يتجاوز على ممتلكات الآخرين.		
14	يقف بعيداً عن الآخرين أثناء اللعب.		
15	ميال للسير والتحرك مع مجموعة من زملائه.		

وكما أسلفنا فإن استخدام هذه القائمة لا يقتصر على تقويم أنماط سلوكية تتعلق بالطالب، بل تستخدم في مجالات عديدة، منها مجال تقويم البرامج، فلو أراد مقوِّم معين تقويم التقرير النهائي لبرنامج معين يمكن تصميم مثل هذه القائمة حتى نوفر فرصة الاستطراد في نقد وتقويم ذلك التقرير، على أن تتضمن القائمة جميع العناصر الأساسية التفصيلية المطلوب توافرها في مثل ذلك التقرير.

وفيما يأتي نموذج بسيط لتقييم تقارير التقويم الذي وضعه رويز وآخرون (Royse et al, 2001)

1- المقدمة: هل المقدمة تقدم فكرة واضحة عن كل من:

أ. مشكلة البحث؟ نعم ( ) لا ( )

ب. طبيعة البرنامج. نعم ( ) لا ( )

جـ. الغرض من التقويم؟ نعم ( ) لا ( )

د. واقعية التقويم؟ نعم ( ) لا ( )

2- مراجعة الأدبيات: هل مراجعة الأدبيات تقدم:

أ. بُنية ذات علاقة بفهم البرامج السابقة والجهود التقويمية الخاصة بتلك البرامج؟ نعم ( ) لا ( )

ب. مسح شامل للأدبيات المعاصرة والسابقة؟ نعم ( ) لا ( )

3- المنهجية: هل المنهجية تصف:

أ. تصميم التقويم؟ نعم ( ) لا ( )

ب. أسلوب اختيار العينة؟ نعم ( ) لا ( )

جـ. خصائص أفراد العينة؟ نعم ( ) لا ( )

د. أسلوب جمع البيانات؟ نعم ( ) لا ( )

هـ. الأدوات المستخدمة؟ نعم ( ) لا ( )

4- النتائج: هل يتضمن هذا الجزء:

أ. نتائج ذات علاقة بموضوع التقويم وأسئلته؟ نعم ( ) لا ( )

ب. جداول لوسائل إحصائية مناسبة؟ نعم ( ) لا ( )

5- المناقشة: هل يوضح هذا الجزء:

أ. تضمينات تطبيقية؟ نعم ( ) لا ( )

ب. تعميمات للنتائج؟ نعم ( ) لا ( )

جـ. تفسيرات مدعمة بالدراسات السابقة؟ نعم ( ) لا ( )

أما طريقة الإجابة على عناصر وفقرات هذه القائمة فيمكن أن تكون بأساليب متنوعة منها وضع اشارة (√) تحت احدى الاستجابتين (نعم، لا) بعد كل فقرة. لتدل على الإيجابية أو السلبية كما ورد في تقرير التقويم. وإذا ما أردنا أن نزاوج بين الاستخدام الكمي والنوعي في عملية تقويم التقرير، فيمكن أن يعقب الإشارة التي نضعها بعد كل فقرة توضيحاً مختصراً أي وصفاً نوعياً لسبب اختيار الإستجابة التي تمثلت باختيار (نعم أو لا)، وبهذه الحالة يمكننا تكوين حكم عام عن تقرير التقويم. وأخيراً فيما يأتي بعض التوصيات الواجب مراعاتها عند تصميم قوائم الشطب في التقويم:

1- حدد بدقة الخاصية أو المهارة المراد تقويمها والعناصر المكونة لها، مستنداً إلى الأدبيات وآراء الباحثين في تلك المهارات والإطلاع على قوائم مماثلة، وخبرتك في الموضوع المقوَّم.
2- أكتب السلوك أو المهارة المراد ملاحظتها وتقويمها بشكل فقرة واضحة محددة يمكن ملاحظتها ببساطة.
3- رتب الفقرات في القائمة بشكل متسلسل يتطابق مع ما يمكن أن يحدث في الواقع، حتى يمكن ملاحظته بشكل واضح وبسيط.
4- حدد الأسلوب الذي على المقوِّم اتباعه في احتساب الدرجة الكلية للسلوك الملاحظ، والمعيار الممكن استخدامه لمقارنة تلك الدرجة به من أجل التوصل إلى الحكم المناسب لذلك السلوك.

مقاييس التقدير Rating - scales

يمكن القول أن ظهور مقاييس التقدير يرجع إلى عدة أسباب لعل أبرزها الذاتية المركزة التي تعكسها أساليب التقرير الذاتي التي كان الباحثون يعتمدون عليها عند جمع المعلومات من الأفراد مباشرة، ومن المعروف أن تحليل استجابات التقرير الذاتي للأفراد تخضع أيضاً للتصورات الذاتية والأطر المرجعية الشخصية للباحثين الذين يقومون بتحليلها، تضاف إلى تلك الصعوبات صعوبة أخرى تلك هي صعوبة تكميم ما يرد في تلك التقارير، ومن تلك المبررات جاء أسلوب مقاييس التقدير كأسلوب يتجاوز إلى حد ما صعوبات السرد في الإجابات الذاتية التقليدية وتفسيرها.

إن مقاييس التقدير تصمم لتحديد نتائج تعلم ومهارات معينة لا يمكن تقويمها من خلال الاختبارات التحصيلية، ومنها على سبيل المثال (النشاطات الرياضية، الأعمال الفنية والمسرحية، الخطابة، الأعمال المختبرية أو تقويم ورقة بحث يكتبها الطالب، أو تقويم بعض جوانب برنامج تربوي ما). وتتميز مقياس التقدير عن قائمة المراجعة أو الشطب، في أن الأخيرة توفر معلومات عن الموضوع المقوَّم أو السلوك الملاحظ من حيث (ظهوره) أو (عدم ظهوره)، بينما يوفر مقياس التقدير مجموعة من الخصائص (أعمال، سلوكيات، نتائج، حوادث... إلخ). يحكم عليها من خلال تدرج وجود الخاصية التي يتم تحديدها مسبقا بشكل كمي.

لقد أشار بعض الباحثين إلى أن مقاييس التقدير تعتبر وسيلة منظمة للحصول على تحديد دقيق للأحكام عن الظاهرة المقاسة، كما أنها توفر معلومات هامة في مجالات عديدة، ويؤكد كرونباك Cronbach إن مقاييس التقدير توفر معيارا جيداً لدراسة نشاطات عدة منها النمو او التطور والتقويم التربوي، أو أداء الطالب في نشاطات متعددة (Alzameli, 1985)

إن استخدام مقاييس التقدير يهيئ للمقوِّم فرصة لملاحظة الأداء المقصود للحكم على فاعليته. وكما في الوسائل التقويمية الأخرى، فإن الغرض الرئيسي لمقاييس التقدير هو تقرير أو تحديد المستوى الذي يظهر الفرد نحو أهداف أو نشاطات معينة، فلو فرضنا أننا نريد تقويم ورقة بحث Research paper مقدمة من أحد الطلبة عن الكتابة في اللغة الانجليزية Writing فعلينا أن نحدد أولاً عناصر القدرة الكتابية وأهمها:

- استخدام الحروف الكبيرة .Capitalization
- أقسام الكلام .Divisions of Speech
- النحو .Grammer
- علامات الترقيم والتنقيط .Puncuation
- معاني المفردات .Vocabulery
- الإملاء .Spelling
- التوثيق والفهرسة .Documentation and References

وبعد الانتهاء من تحديد عناصر السمة المطلوب تقويمها نبدأ بتصميم فقرات المقياس، ولابد لنا أن نقرر ما إذا كانت جميع العناصر المحددة للكتابة يمكن أن تقوّم بنفس المستوى والأهمية، أم أن بعضها يعطى أهمية أكبر من العناصر الأخرى، أي هل ستخصص له درجات أكثر من الدرجات المخصصة للعناصر الأخرى؟ ففي المثال السابق، ونظراً لأن الجهد الذي يبذله الطالب في مجال التوثيق والفهرسة هو أكبر من الجهد المبذول في الجوانب الأخرى، فلا بأس أن تكون الدرجة القصوى لهذا العنصر أكبر من الدرجات القصوى التي تخصص للعناصر الأخرى، كما أن هناك أمراً آخر لابد من تحديده قبل تطبيق المقياس، ألا وهو عدد الاستجابات (مستويات التقدير) التي توضع بعد كل فقرة من فقرات المقياس، هل هي (ثلاثة، خمسة، سبعة)، علماً بأن الشائع في هذا المجال هو المقياس الخماسي، أو أي مقياس ذو أبعاد فردية حتى تكون هناك نقطة وسطية واضحة ومحددة، فمثلاً المقياس الذي ينتهي بخمسة تقديرات (1، 2، 3، 4، 5) يكون التقدير (3) ممثلاً لمنتصف المسافة في السمة المراد قياسها من خلال فقرات ذلك المقياس.

شكل (5-3) مقياس لتقدير مستوى ورقة بحث في مادة اللغة الإنجليزية

العنصر المقوّم	التقديرات					الملاحظات
	1	2	3	4	5	
النحو						
التنقيط						
المصطلحات						
التهجأة						
استخدام الحروف الكبيرة						
أقسام الكلام						
التوثيق والفهرسة						
الدرجة الكلية						

ومن الجدير ذكره في هذا المجال أن استخدام مقاييس التقدير لا ينحصر

في قياس وتقويم مهارات الطالب فحسب، بل إنه ضروري وعملي عندما يستخدم من قبل مقومي البرامج حيث يحدد المقوم الجوانب الهامة في البرنامج المراد تقويمه ويضعها بشكل فقرات تتبعها مستويات التقدير على وفق أهمية كل فقرة وفي ضوء معايير محددة سلفاً يستطيع إصدار أحكامه عن ذلك الجانب الخاص بالبرنامج المقوَّم.

أنواع مقاييس التقدير:

على الرغم من وجود أنواع عديدة لمقاييس التقدير، لكن الأنواع الثلاثة الآتية هي الأكثر شيوعاً واستخداماً: (Alzameli, 1985)

1- مقياس التقدير العددي Numerical rating - Scale

وهو من أبسط الأنواع حيث يتطلب من المقوّم أن يؤشر أو يضع دائرة حول الرقم الذي يمثل درجة تقويم بالأداء المقوّم، وفي هذا النوع من الضروري إخبار مستخدم المقياس بأن الرقم الأكبر هو الذي يمثل الحد الأعلى للأداء.

مثال:

- مدى وضوح أهداف البرنامج لدى الموظف المسؤول عن شؤون الموظفين في المؤسسة؟

(1) (2) (3) (4) (5)

2- مقياس التقدير البياني: Graphic rating - Scale

وفيه تتبع الخصائص المقوَّمة مصطلحات تمثل الاستجابات المطلوب اختيار واحدة منها:

مثال:

مدى مشاركة المستفيدين من البرنامج في عملية صنع القرارات.

لا يشاركون أبداً	نادراً	أحياناً	غالباً	دائماً
( )	( )	( )	( )	( )

3- مقياس التقدير الوصفي Discriptive rating- scale

وفي هذا النوع يتم استخدام عبارات وصفية لتوضيح مستويات الأداء المطلوب تقويمه.

مثال:

إلى اي مدى يشارك هذا الطالب (س) في المناقشات الصفية؟

لا يشارك إطلاقاً.	( )
يشارك مثل الآخرين في الصف.	( )
يشارك أكثر من أي طالب في الصف.	( )

ويمكن وضع أرقام تحت مستويات التقدير الموضحة في النوعين الثاني والثالث من مقاييس التقدير السابقة الذكر.

خطوات إعداد مقاييس التقدير:

فيما يلي مجموعة من الإرشادات التي يمكن لمصمم مقياس التقدير الانتباه لها من أجل الحصول على مقياس صادق وموضوعي يحقق الغاية التي يستخدم لأجلها:

1- ضرورة تحديد الأهداف التي سيستخدم المقياس لتحقيقها، وأن تكون شاخصة في ذهن الباحث في كل خطوة من الخطوات الأخرى.

2- تحديد العناصر المكونة للخاصية المراد ملاحظتها وتقديرها في ضوء الأهداف التي تمّ تحديدها في الخطوة السابقة والأدبيات والدراسات السابقة التي تناولت نفس الموضوع.

3- البدء بصياغة فقرات تمثل العناصر الممثلة للأداء أو المهارة المطلوب تقويمها، على أن تصاغ بشكل واضح ودقيق وتكون توصيفاً للسلوك المراد ملاحظته.

4- تحديد الوقت المناسب لملاحظة السلوك المطلوب، والتأكد من أهميته لتزويد الباحث بالبيانات الكافية التي يهدف إلى جمعها.

5- تحديد التقديرات الخاصة بمستويات الأداء بشكل واضح سواء كان بالأسلوب الكمي، أو بالأسلوب الوصفي النوعي.
6- عدم الإقلال من نقاط التقديرات وعدم الإكثار منها أيضاً، لأن العدد القليل يؤدي إلى قياس غير دقيق، أما الإكثار من التقديرات فيؤدي إلى صعوبة التمييز بين تلك التقديرات ويتطلب وقتاً أطول لتحديد التقدير المناسب للسلوك المقوّم.
7- عرض المقياس على مجموعة من المتخصصين وذوي الخبرة وأخذ ملاحظاتهم بنظر الاعتبار في الصياغة النهائية للمقياس قبل تطبيقه.
8- التأكد من صدق المقياس وثباته وفقاً للأساليب المعروفة للتحقق من الخصائص السيكومترية لأدوات القياس.
9- ضرورة اختيار الأفراد الذين يتسمون بالموضوعية والدقة وتدريبهم بشكل كافي قبل البدء بعملية الملاحظة والتقدير.

بعض المشكلات المصاحبة لاستخدام مقاييس التقدير:

على الرغم من أهمية مقاييس التقدير في توفير معلومات كمية أو نوعية مفيدة للمقوّم عن الشيء المراد تقويمه، فإن هناك العديد من نقاط الضعف التي يمكن أن يحملها هذا النوع من المقاييس نرى من الضروري أن نطرحها أمام القارىء لأجل أن يتجنبها الباحثون خلال استخدامهم لها.

إن المشاكل المصاحبة لاستخدام هذه المقاييس يمكن أن تكون على نوعين. يتمثل النوع الأول في وجود عدد من العوامل التي تحد من رغبة المقوّم في إعطاء تقويم صادق ودقيق وفق التعليمات الخاصة باستخدام المقياس.

أما النوع الثاني فيتمثل بوجود عوامل تقلل من قدرته على إعطاء تقديرات ثابتة وصحيحة، من مشكلات النوع الأول ما يأتي:
1- قد لا يرغب القائم بعملية التقويم في تحمل الأعباء التي يتطلبها منه القيام بعملية التقويم.

2- أن المقوّم قد يكون على علاقة مع الشخص المراد تقويمه بحيث يريد تحاشي ذكر الجوانب السلبية ونقاط الضعف في أدائه.

أما المشكلات من النوع الثاني اي التي تتعلق بوجود عوامل تقلل من قدرة المقوّم على إعطاء تقديرات ثابتة وصحيحة فيمكن ذكر أهمها بالآتي:

1- محدودية الفرص المتاحة أمام المقوّم لملاحظة سلوك الفرد المراد تقويمه، فقد يحدد الباحث ساعة واحدة لملاحظة السلوك العدواني لخمسة طلاب خلال تواجدهم في ساحة المدرسة أو في الصف، لكنه يفاجأ بعدم وجود إثنين منهم لتغيبهم عن المدرسة ذلك اليوم، أو قد تنتهي الساعة دون أن يستطيع ملاحظة إثنين منهم، وفي أحيان أخرى قد تنتهي الساعة دون أن يظهر اي من الطلاب الخمسة اي نمط من أنماط السلوك العدواني المراد ملاحظته.

2- صعوبة تقويم بعض الأنماط السلوكية غير الظاهرة بواسطة مقياس التقدير واقتصاره على تقدير الأنماط الظاهرة اي التي تقع تحت ملاحظة المقوّم، مما يجعل منه أداة غير متكاملة لقياس جميع أنماط السلوك الإنساني:

3- غموض معنى الفقرة التي يراد تقويم أداء الفرد وفقاً لها، مما يجعل المقوم في حيرة من أمره في تفسيرها وتحديد مدلولاتها لكي يتوصل إلى قناعة في تحديد الدرجة المناسبة لها، فلو عرضت على مقوّم ما عبارة (الشعور بالمواطنة) وهو يقوّم سلوك طفل في مدرسة ابتدائية مثلاً، فما المقصود بهذه العبارة الواسعة، وما هي حدود الأداء الذي يدل عليها لكي يستطيع المقوِّم تحديد الدرجة التي يضعها له؟ هل تعني عدم الكتابة على الجدران؟ أو عدم رمي الأوساخ على الأرض ووضعها في سلة النفايات؟ أم مساعدة المعلم في تهيئة وتنظيف غرفة الصف؟ هذه احتمالات محددة من مؤشرات أخرى كثيرة تدل على الشعور بالمواطنة، فهل يستطيع المقوِّم إذا ما لاحظها تحديد مستوى الطالب على هذه الفقرة؟ إنها بلا شك عملية صعبة وتحتاج إلى وقوف وملاحظات متعددة مما لا يتاح له إنجازها في

وقت ضيق محدد أمام المقوّم، مما يدفعه إلى وضع التقدير بناءً على تصوراته وانطباعاته التي قد تكون عامة عن ذلك الطالب مما لا يحقق الدقة في التقدير.

4- إهمال بعض المقوّمين لوضع مرجع معياري منتظم يمكن الحكم إستناداً له. إذ من الملاحظ أن الكثير من مقاييس التقدير تتطلب إعطاء أحكام عن الأفراد الذين يجري تقديرهم من خلال استخدام مجموعة من الفئات التي يفترض اختيار واحداً منها تنطبق عليه الحالة المقوّمة، ففي حالة عرض مستويات التقدير على المقومين: (متفوق، ممتاز، جيد جداً، جيد، مقبول، غير مقبول) والاستفسار من المقوم عن المستوى المحدد للتقويم؟ أي هل الشخص الذي يعطى (جيد) هو من بين أفضل عشرة أشخاص في المجموعة التي يقارن بها؟ أو في الربع الأعلى؟ أو في النصف الأعلى؟ أو أنه فقط ليس من بين أدنى عشرة في الصف؟ وما طبيعة المجموعة التي يفترض أن يقارن معها؟ وأسئلة أخرى كثيرة يمكن طرحها في هذا المجال، فكيف يفترض أن يحدد المقوّم مستوى القدرة على إعطاء الأحكام عن شخص في وظيفة ما بمستوىً معيناً من الخبرة، لذلك فإن المقوِّم الذي يسعى أن تكون تقديراته دقيقة وموضوعية يفترض أن يحدد معايير مفترضة لكل عنصر من العناصر المراد تقديرها عند التقويم، وأن الاختلاف في تفسير الفقرات أو الاختلاف في تعريف المجموعة المرجعية، واختلاف الخبرة في ذلك المعيار المرجعي كلها تؤدي إلى التباين من مقوّم لآخر بما يؤدي إلى اختلاف التقديرات، وهذه الظاهرة كثيراً ما تلاحظ في تقديرات المعلمين لطلابهم في مهارات تعليمية متنوعة.

5- ميل العديد من المقومين إلى إعطاء تقديرات تتراوح من الوسط إلى التقديرات الأعلى وقلة التقديرات التي تنحصر تحت المتوسط، وهذا ما لوحظ بشكل أكثر تكراراً عند تقويم المدراء لموظفي مؤسساتهم تعاطفاً معهم، وتحاشياً للإضرار بزملائهم، (ثورندايك وهيجن، 1989).

6- أثر الهالة: إن محدودية خبرة المقوّم في الشخص الذي يجري تقويمه، ونقص الفرص المتاحة لمشاهدة صفات محددة واردة في مقياس التقدير، وتأثير التحيزات الشخصية التي تؤثر على مستوى الميل العام للشخص تتضافر جميعها لتحدث نوعاً آخر من عدم الدقة والموضوعية، ذلك هو الميل إلى إعطاء تقديرات نتيجة انطباعات عامة دون التفريق بين العناصر المحددة وهذا ما يسمى بتأثير هالة الفرد Halo effect على عملية التقدير. فمثلاً في حالة تقدير المستوى العام لأداء الطالب وفقاً لمعايير محددة في مقياس التقدير، نلاحظ ميل المعلم لإعطاء الطالب مستوى واحداً لجميع المعايير انطلاقاً من إنطباعاته السابقة عن ذلك الطالب سلبية كانت أم ايجابية دون النظر إلى كل عنصر من عناصر الأداء على حده حسب مستوى الطالب وفقاً لذلك العنصر.

الأثنوغرافيا: Ethnography

تعتبر الإثنوغرافيا أسلوباً من أساليب البحث والتقويم لا سيما التقويم النوعي، وهي وصف تحليلي للمواقف الاجتماعية بما فيها من الأفراد والجماعات والتركيز على ملاحظة الجوانب الثقافية والمعتقدات والعادات والتقاليد التي يمارسوها وجمع البيانات التي تؤدي للحصول على تصوراتهم وسلوكياتهم في البيئات الاجتماعية التي يعيشون فيها. (أبو زينة وآخرون، 2005).

ويؤكد بعض الباحثين على أن الأثنوغرافيا دراسة مركّزة لخصائص ثقافة معينة وأنماط تلك الخصائص، وإذا ما طُبقَ هذا الأسلوب بشكل جيد، فإن قارىء التقرير النهائي للدراسة الأثنوغرافية سيكون قادراً على فهم أبعاد تلك الثقافة بشكل وافي وكأنه خبرها من قبل. لقد شاع هذا الأسلوب بين علماء الاجتماع والأنثروبولوجيا واستخدموه بكثرة في دراساتهم عبر العقود الماضية لا سيما في القرن العشرين حينما ركزوا على دراسة الثقافات أو الحضارات لجماعات معينة والاهتمام بوصف الحياة التي تعيشها المجتمعات وتوثيق المشاهدات التي تلاحظ من قبلهم مباشرة وبمشاركتهم للحياة الاجتماعية لتلك

المجتمعات ثم تحليل البيانات التي يتم جمعها من الميدان تحليلاً وصفياً اي بأسلوب نوعي.

لقد انتقل استخدام الأسلوب الأثنوغرافي من الأنثروبولوجيا الثقافية الاجتماعية إلى الميدان التربوي منذ أواخر القرن التاسع عشر وأوائل القرن العشرين، عندما ركز بعض الباحثين التربويين دراساتهم على الثقافات البدائية Primitive cultures من خلال زياراتهم لمجتمعات أخرى مغمورة لفترة من الزمن، ولاحظوا تلك المجتمعات عن قرب، ثم وثقوا ما رأوه وما سمعوه عن كثب، ثم قورنت حياة تلك المجتمعات بحياة المجتمع الأمريكي، ولعل أبسط مثال على تلك الدراسات ما قامت به الباحثة الشهيرة مارجريت ميد (Marigret Mead, 1928) التي ركزت دراساتها على الأطفال والمراهقين وتأثير الحضارة على الشخصية في مجتمع (ساموا Somoa) الواقعة في قارة أفريقيا، لقد استخدم الباحثون في دراساتهم للجوانب الثقافية التربوية في تلك المجتمعات الملاحظات والمقابلات كوسيلة معيارية لجمع البيانات من الميدان، ولا بد من أن نتذكر الدراسات الأثنوغرافية التي قام بها بعض الباحثين من علماء التربية وعلم الاجتماع في جامعة شيكاغو الأمريكية في العشرينات وحتى الخمسينات من القرن الماضي واستخدموا فيها أسلوب دراسة الحالة الواحدة single case سواء كانت تلك الحالة فرداً واحداً أو مجموعة أو منطقة سكنية بأكملها. وكمثال على تلك الدراسات دراسة حياة المهاجرين البولنديين في الولايات المتحدة والتي أجريت عام 1927 والتي زودت الباحثين بتصورات ذاتية مفصلة لما يشعر به أفراد تلك الجالية من حياة هامشية كانوا يحسونها وهم يعيشون في الولايات المتحدة الأمريكية .(Creswell, 2002).

ولقد إزدادت حركة البحث الإثنوغرافي في المجال التربوي بعد الخمسينات وتصاعدت في تطورها في الثمانينات من القرن المنصرم، وأصبحت تركز على موضوعات مهمة من مثل التركيب العرقي والثقافي، والطقوس التي تتميز بها الصفوف الدراسية في المراحل الابتدائية والثانوية،

وفحص عملية صنع القرارات التربوية ومحتوى المناهج والتدريس. وأصبح علماء الانثروبولوجيا التربوية يركزون دراساتهم على مقاطعات ثقافية معينة من الولايات المتحدة لدراسة تاريخ الحياة اليومية المهنية وتحليل دور الفرد فيها، كما أجريت دراسات إثنوغرافية محدودة من داخل الصفوف الدراسية باعتبارها مجتمعات مصغرة.

إن البحث الإثنوغرافي يقوم على مجموعة من الافتراضات منها :

1- إن سلوك الأفراد يتأثر كثيراً بالبيئة التي يعيشون فيها، ولفهم ذلك السلوك لا بد من فهم تلك البيئة .
2- إن البيانات التي يرتكز عليها البحث الإثنوغرافي تجمع من موقعها الحقيقي مباشرةً.
3- إن عملية تفسير البيانات تجري في إطار البيئة التي جمعت منها.
4- لا يهتم الأسلوب الإثنوغرافي بتعميم النتائج قدر اهتمامه بالوصف الدقيق الكافي للموقف المدروس.
5- يهتم الباحث بالمحافظة على الموضوعية ويبعد تحيزاته الذاتية والقيمية عن المواقف المدروسة.
6- تُعتبر تصورات أو ادراكات الأفراد للمواقف هي الموجه الأساسي لسلوكهم نحو الآخرين.
7- يهتم الأسلوب الاثنوغرافي بالتفاعل بين الباحث والمشاركين أي المبحوثين في الموقف المدروس.
8- يتصف الباحث الأثنوغرافي بحساسيته وانتباهه الزائد لكل عنصر من عناصر الموقف المدروس، دون تأثيره على متغيرات ذلك الموقف، وتدوين أحداثه بدقة وأمانة.

ويُشير جول وآخرون (Gall, et al, 2003) إلى أن الأسلوب الأثنوغرافي يتميز بثلاثة خصائص أساسية هي:

1- التركيز على تاثير النماذج الثقافية على السلوك الإنساني حيث يدرس الباحث الأثنوغرافي الأفراد لتحديد كيف أن سلوكهم وقيمهم ومعتقداتهم وعاداتهم ومحرماتهم هي نماذج ثقافتهم التي يعيشون بظلها، لذلك ينظر الباحث لكل فرد على أنه وثيقة تزوده بمعلومات هامة عن المجتمع الأكبر.

2- يركز هذا الأسلوب على تصورات الأفراد وإدراكهم للثقافة التي يعيشون فيها، أي كيف يحددون الواقع والأحداث والخبرات، وبالتالي وصف الثقافة كما يراها أعضاء المجتمع.

3- يركز الأسلوب على دراسة المحيط الطبيعي الذي تظهر فيه الثقافة كما هي لذلك يتجنب هذا الأسلوب تقديم أي ظاهرة مصطنعة في المحيط، بل إنه يركز كل اهتمامه على جميع الجوانب التي تكشف الأنماط الثقافية في ذلك المحيط.

أسلوب جمع البيانات وفقاً للأثنوغرافيا:

يبدأ الاثنوغرافيون دراستهم لحضارةٍ أو ثقافة معينة بوضع افتراضات وتصورات تصاغ باسلوب استقرائي يعتمد على جمع الدلائل والشواهد المختلفة والوصول منها الى أفكار مجردة ومحددة وفقاً للطرق الآتية:

1- جمع الادلة العلمية من الميدان مباشرةً عن طريق اختلاطهم ومشاركتهم في الحياة الاجتماعية للافراد أو الاحداث وكما هي في سياقها الطبيعي دون تدخلٍ أو تغيير من قبل الباحثين.

2- اجراء مقابلات شخصية مع افراد من المجتمع المدروس للوصول الى وجهات نظر تفسر الكثير من الظواهر والبناءات الاجتماعية التي يريد الباحث فهمها وسبر أغوارها.

3- يقوم الباحث الاثنوغرافي بتوثيق وجهات نظر الافراد الذين يلاحظهم أو يقابلهم أو يعايشهم أو يختلط بهم أو من خلال استخدام أية اساليب

بحث اخرى ويُفضل استخدام وسائلٍ متعددة دون الاعتماد على وسيلة واحدةً لجمع المعلومات.

4- يتم تحليل البيانات التي قام الباحث بجمعها بأساليب علمية يغلب عليها الاسلوب النوعي التحليلي للوصول الى استنتاجات علمية يمكن ان توفر أرضية نظرية ينطلق منها باحثون آخرون لدراسات لاحقة، وقد يتحول البعض الآخر منها الى نظريات علمية تصلح لتفسير الكثير من الاعراف والسلوكيات الثقافية والاجتماعية والتربوية.

الاستبيان Questionnaire

يعتبر الاستبيان من أكثر أدوات البحث والتقويم شيوعا واستخداما في جمع معلومات تتعلق بموضوعات مختلفة، كتحديد الاتجاهات والمشكلات، أو معرفة الأسباب التي تؤدي إلى حدوث ظاهرة معينة وغير ذلك.

ويشيع استخدام الاستبيان في مجال التقويم التربوي على وجه الخصوص للتعرف على آراء المستفيدين من برنامج معين، أو العاملين فيه عن الصعوبات أو المشكلات التي يواجهها ذلك البرنامج، أو التوصل إلى أحكام عامة تتعلق بفاعليته أو نقاط قصوره. ولعل سهولة استخدام الاستبيان هي التي أدت الى انتشاره بشكل أكبر من انتشار أدوات التقويم الأخرى.

ولاستخدام الاستبيان ظروف ومتطلبات على المقوم أن ينتبه لها، ومنها توحيد ظروف التطبيق على أفراد العينة بما يساعد في الحصول على معلومات موضوعية تساعد على إجراء مقارنات فيما بينها، كما يتطلب أيضا أن يكون الهدف المطلوب تحقيقه من الاستبيان واضحا ومحددا للمستجيبين، وأن تكون المتغيرات التي يتضمنها ذات فائدة ويتم استخدامها في التحليل للتوصل إلى المعلومات المطلوبة التي توفر أرضية صلبة تساعد متخذ القرار على أن يختار القرارات الصائبة حول الموضوع قيد التقويم.

والاستبيانات المستخدمة في التقويم قد تكون جاهزة يقتبسها المقوم من

بحوث أخرى سابقة ويعمل على تكييفها وفق طبيعة الهدف من التقويم الذي يقوم به، أو قد تكون من عمل المقوم نفسه. لكن الأمر الهام في هذا المجال أن يخضع الاستبيان في كلا الحالتين إلى توافر المؤشرات السيكومترية الخاصة بصدقه وثباته ومؤشرات ضرورية أخرى. فضلاً عن ذلك فإن من الضروري أن يتم إخراج الاستبيان بشكل جذاب يشجع المستجيب على متابعة فقراته حتى النهاية ويعطيه اهتماماً أكثر، ولعل الرسالة المرفقة بالاستبيان والمصاغة بشكل رصين وعبارات جميلة ستدفعه إلى الحرص على أن يدقق بإجاباته قبل أن يضعها على الاستبيان.

علاوة على ما مرَ ذكره فإن الاستبيان يصلح في مجال التقويم لتوفير معلومات مهمة حول:

1- تحديد اتجاهات الأفراد نحو المشروعات والبرامج التربوية. مثل تحديد اتجاهات المعلمين نحو استخدام التقويم الذاتي المستخدم في المدارس.
2- تحديد الصعوبات والمعوقات التي يواجهها تطبيق مشروع تربوي معين.
3- التعرف على تقويمات المعلمين والطلبة للمناهج والكتب الدراسية المستخدمة في مرحلة تعليمية معينة.
4- التعرف على آراء العاملين في سوق العمل عن جودة مخرجات مؤسسة تعليمية معينة، أو مدى ملائمة برامجها لطبيعة العمل الذي سيزاوله أولئك الخريجون في المستقبل.
5- تقويم المقررات الدراسية الجامعية من قبل الطلبة.
6- تقويم أداء المعلم سواء من قبل الطلبة، أو الزملاء، أو الإدارة المدرسية.

وهناك موضوعات أخرى كثيرة يستخدم فيها الاستبيان كأداة رئيسية في عملية التقويم.

مفهوم الاستبيان

ورد الاستبيان في أدبيات مناهج البحث والتقويم، على أنه استبيان (Questionaire) وعلى أنه مسح Survey وفي كلا الحالتين (الاستبيان والمسح)

يمثلان مصطلحين للمفهوم نفسه. وفيما يلي عدد من تعاريف الاستبيان :

1- استمارة مطبوعة تحتوي على مجموعة من الأسئلة المنظمة بعضها مع بعض، تستخدم غالبا في الدراسات الوصفية.

2- "أداة لتجمع بيانات ذات صلة بمشكلة بحثية معينة عن طريق ما يقرره المستجيبون لفظياً في إجاباتهم عن الاسئلة التي يتضمنها الاستبيان" (عطفية، 2002، ص 286).

3- "مجموعة مكتوبة من أسئلة التقرير الذاتي Self-report تجيب عنها عينة البحث" (Gay, Mills & Airasian, 2006, P. 163)

4- "نموذج يتضمن مجموعة من الأسئلة الشخصية أو الديموغرافية، يجيب عنها عينة من المستجيبين بوضع إشارة في الاختيار المناسب" (Creswell, 2002, P. 648)

أنواع الاستبيان

الاستبيان على ثلاثة أنواع، هي: الاستبيان المفتوح، والاستبيان المغلق، والاستبيان المغلق- المفتوح. وفيما يأتي عرض لكل نوع من الأنواع الثلاثة.

1 - الاستبيان المفتوح

هذا النوع من الاستبيان هو الأكثر انتشارا، ويتميز بترك الحرية للمفحوص في التعبير عن آرائه بالتفصيل، مما يساعد الباحث في التعرف على الأسباب والعوامل والدوافع التي تؤثر على الآراء والحقائق، ولكن من مآخذ هذا النوع من الاستبيان أن المبحوث قد يجيب بطريقة تختلف عن قصد الباحث، وتدني نسبة الردود على هذا النوع من الاستبيان، بالإضافة إلى أن المفحوصين لا يتحمسون عادة للكتابة عن آرائهم بشكل مفصل، ولا يمتلكون الوقت الكافي للإجابة عن اسئلة تتطلب منهم جهدا، كما أن الباحث يجد صعوبة في دراسة وتصنيف وتحليل إجابات المفحوصين بشكل يساعده للإفادة منها، وإن تفريغ المعلومات من هذا النوع من الاستبيان أمر عسير إحصائياً.

ومن أمثلة الاستبيان المفتوح سؤال الطالب عن رأيه بالمقرر الدراسي، أو طريقة التدريس.

مثال لاستبيان مفتوح

س1- ما أكثر ما أعجبك عند دراسة هذا المقرر؟

س2- ما أكثر ما ضايقك عند دراسة هذا المقرر؟

س3- أذكر أية مقترحات أو ملاحظات تعتقد أنها مهمة في تطوير تدريس المقرر.

2 - الاستبيان المغلق

هذا النوع من الاستبيان يتضمن عددا من الأسئلة المقيدة، وإزاء كل سؤال عدد من الإجابات (البدائل) أقلها اثنين، وعلى المستجيب أن يختار من بين تلك الاجابات إجابة واحدة أو أكثر وذلك وفقا للتعليمات الواردة في الاستبيان. وتتم صياغة الأسئلة وطريقة الإجابة على تلك الأسئلة بعدة طرق منها:

1- تقديم السؤال ومعه إجابتان بديلتان مثل: نعم أو لا، وعلى المستجيب أن يختار إجابة واحدة من بين الإجابتين، ويضع علامة أمام الإجابة التي اختارها.

مثال:

هل تفضل العمل بمهنة التدريس؟ نعم ( ) لا ( )

2- تقديم السؤال ومعه أكثر من إجابتين بديلتين، وعلى المستجيب أن يختار إجابة واحدة من بين الإجابات المقدمة.

مثال لاستبيان مغلق:

بعد كل عبارة مما يأتي اختر التقدير المناسب، وذلك بكتابة أحد الأرقام من (1-5):

لا أوافق تماما	لا أوافق	محايد	أوافق	أوافق تماما
1	2	3	4	5

العبارات	التقدير
حدد الأستاذ الأهداف والمفردات والأنشطة بشكل واضح ومفهوم في أول محاضرة.	
شرح الأستاذ للمحاضرة كان واضحا وسهلا.	
كان الأستاذ متعاونا معنا.	
شجعني الأستاذ على التفكير أكثر من مجرد قبول الحقائق وحفظها.	
استطاع الاستاذ أن يثير اهتمامي بهذا المقرر.	
شجعني الاستاذ على طرح الاسئلة والمناقشة.	
استخدم الأستاذ الوسائل التعليمية المناسبة عند تدريس المقرر.	
الامتحان كان شاملا للمادة التي درستها.	
لقد تعلمت الكثير من دراسة هذا المقرر.	
بصورة عامه كان أستاذ هذا المقرر أستاذا جيدا.	

3- تقديم السؤال ومعه أكثر من إجابتين يختار المستجيب واحدة أو أكثر منها دون تفضيل لإجابة من الإجابات التي اختارها على الإجابات الأخرى، أو ترتيب مجموعة من العبارات وفقا لأهميتها، ويمكن للباحث أن يحدد للمفحوص عددا معينا من الإجابات ليختار من بينها.

مثال:

رتب العبارات التالية في ضوء أهميتها بالنسبة لك:

- أستخدم الإنترنت كطريقة تدريس كونها وسيلة تكنولوجية حديثة ( ).
- أستخدم الإنترنت كونها وسيلة فعالة في التدريس ( ).
- أستخدم الإنترنت لأنها تحفز التفكير لدى الطلاب ( ).
- أستخدم الإنترنت لتسهيل العملية التعليمية ( ).

3 - الاستبيان المغلق- المفتوح

إن هذا النوع من الاستبيان يحاول أن يجمع بين خصائص النوعين السابقين، ويتغلب في نفس الوقت على عيوبهما، ويتكون من أسئلة مغلقة، وأسئلة مفتوحة، تعطي الحرية للمفحوص في الإجابة عن عناصر يبحث الباحث عن إحابات لها، ويتعين على الباحث أن يلم بطرائق الإحصاء المناسبة في مثل هذه الحالات.

مثال:

هل تؤيد الاكتشاف الموجه كطريقة تدريسية؟

نعم ( ) لا ( )

إذا كانت الإجابة (لا) فما هي الأسباب؟

-

-

-

خطوات تصميم الاستبيان

يمر إعداد الاستبيان بخطوات علمية محددة ومتفق عليها، يتعين على الباحث مراعاتها عند تصميمه استبيانا جديداً، وهذه الخطوات هي:

(1) إعداد أسئلة الاستبيان

- الدراسة الاستطلاعية:

يتم فيها تطبيق استبيان استطلاعي، أو إجراء مقابلة مع عينة صغيرة من الفئة المستهدفة في البحث، وذلك بهدف جمع معلومات أولية عن موضوع البحث. فإذا أراد الباحث أن يقوّم برنامجاً دراسياً معيناً، يمكن أن يسأل الطلبة أو الخريجين السؤال الآتي:

س/ ما هي أبرز نقاط القوة والضعف في البرنامج الذي درسته؟

● مراجعة الاستبيانات السابقة:

تفيد هذه الخطوة في الاستفادة من الفقرات الواردة في الاستبيانات السابقة، وتكوين فكرة عن عدد الفقرات، ونوع البدائل المستخدمة، وما إلى ذلك.

(2) صفحة التعليمات

وتسمى كذلك صفحة التصدير Cover Letter، وهي بوابة يدخل من خلالها المستجيب إلى الاستبيان، فإذا أحسن إعدادها أدت الهدف المطلوب، وشجعت المستجيب على المشاركة في البحث، أما إذا لم تكن مصاغة بشكل جيد، أو لم تتضمن المعلومات المطلوبة فإنها ستؤدي دورا سلبيا لما سيكوِّنه المستجيب من اتجاهات قبل قراءة أسئلة الاستبيان.

ينبغي أن تتضمن صفحة التعليمات:

1- مقدمة تحتوي على عنوان البحث وأهدافه.
2- وصف كمي ونوعي لأسئلة الاستبيان.
3- طريقة الإجابة عن الاستبيان والوقت المتوقع لانجازه.
4- خاتمة تشكر فيها المستجيب، مع ذكر طريقة يمكنه من خلالها التواصل مع الباحث إذا أراد استفساراً عن شيء ما (بريد الكتروني)، أو (رقم هاتف).

(3) إعداد الصورة الأولية للاستبيان:

تشتمل الصورة الأولية للاستبيان على صفحة التصدير، وأسئلة الاستبيان items، ومقياس التقدير rating scale، والمتغيرات الديموغرافية ذات الصلة بموضوع البحث.

صدق وثبات الاستبيان:

في الغالب يلجأ الباحثون إلى المحكمين في تحديد الصدق، ويسمى بالصدق الظاهري، وأما الثبات فيتم حسابه من خلال تطبيق الاستبيان على عينة استطلاعية من الأفراد من غير عينة البحث الأساسية، ومن أسهل الطرق وأكثرها شيوعا حساب معامل ألفا - كرونباخ بواسطة برنامج SPSS.

قواعد أساسية في صياغة بنود الاستبيان

1- تجنب الصياغة التي تُحرِج المجيب، وذلك حتى لا يرفض الإجابة أو يتعمد تقديم إجابة خاطئة .

2- الوضوح والدقة في الصياغة، وذلك بتجنب الكلمات التي قد لا يتفق على مدلولها الباحث والمجيب مثال ذلك: (غالبا، كثيرا، . . . . .الخ).

3- عدم الإطالة في السؤال؛ فكلما كان السؤال قصيرا كلما كان ذلك دافعاقويا للإجابة عليه.

4- تجنب الأسئلة المركبة وهي التي تحتوي على فكرتين، وكذلك الأسئلة التي تشتمل على (واو) العطف، كأن يكون السؤال (هل ترى أن للراتب والمنزلة الاجتماعية أثرا على الإقبال على مهنة التدريس؟).

5- أن تكون الأسئلة محددة بحيث يمكن للمجيب معرفة المطلوب تماما. مثال: هل تعتقد أن النشاط اللاصفي يساعد على تحقيق الأهداف التربوية؟

6- ألا تكون الأسئلة قابلة للتأويل.

7- ألا تحتاج الأسئلة إلى عمق في التفكير، فقد يستطيع المجيب الإجابة عليها وقد لا يستطيع.

8- ضرورة اختيار الكلمات التي يعرف المجيب معناها عند صياغة الأسئلة، فكثيرا ما يؤدي الاستبيان إلى نتائج مضللة بسبب عدم اختيار الكلمات المناسبة.

9- حاول قدر الإمكان أن لا تجعل الاستبيان طويلا جدا يؤدي إلى تعب المستجيب والى شعوره بالملل.

مزايا الاستبيان

1- يؤدي الى اختصار الوقت والجهد في جمع البيانات، بخاصة إذا تم إرساله بالبريد.

2- من الممكن تغطية أماكن متباعدة في أقصر فترة ممكنة.

3- يعطي للمبحوث الحرية في اختيار الوقت المناسب لتعبئة الاستبيان وحرية التفكير.

4- قد يقلل من التحيز سواء من قبل الباحث أو المبحوث، فبالنسبة للباحث يتم طرح الأسئلة نفسها على جميع المبحوثين وبالأسلوب نفسه. أما المبحوث فانه يبدي رأيه بصدق وصراحة دون خوف نظرا لأنه في كثير من الأحيان لا يكتب اسمه على الاستبيان.

5- لا يحتاج إلى عدد كبير من مساعدي الباحث لتوزيعه على المبحوثين.

عيوب الاستبيان

1- الغموض في صياغة الأسئلة يؤدي إلى عدم فهم المفحوص ماهو مطلوب منه، وبذلك ترتفع نسبة الخطأ في الإجابة، أو ترك أسئلة بدون إجابة .

2- تفضيل الباحث لنوع معين من الأسئلة يجعله يتحيز لا شعوريا إلى هذا النوع، وبالتالي يمكن أن تتكّون الأسئلة بطريقة قد لا تكون في مصلحة الاستبيان.

3- عدم قدرة الباحث على معرفة بعض الأمور الانفعالية أو العاطفية من قبل المبحوث أثناء الإجابة.

4- لا يصلح إلا مع المبحوثين الذين يجيدون القراءة والكتابة.

5- صعوبة الأسئلة أحيانا، مما يتطلب قدرا كبيرا من الشرح والتوضيح.

6- زيادة عدد الأسئلة مما يؤدي إلى ملل المفحوصين وعدم تجاوبهم مع الاستمارة.

7- قد تتأثر إجابات المفحوص بالآراء المختلفة للآخرين وبذا تكون إجاباته غير معبرة عن رأيه الشخصي.

8- يؤدي التعب والملل إلى إجابات رتيبة لا تعبر بصدق عما يحسه المستجيب.

نموذج مختصر لاستبيان خاص بتقويم نظام تطوير الأداء المدرسي في مدارس سلطنة عمان (الزاملي وآخرون، 2008)

أولاً - معلومات عامة

- المنطقة التعليمية: ....
- الوظيفة:

1- رئيس قسم ( )
2- مشرف تربوي ( )
3- موجه إداري ( )
4- مدير مدرسة ( )
5- مساعد مدير مدرسة ( )
6- معلم أول ( )

- سنوات الخبرة المهنية: ....
- المؤهل الدراسي:

1- دبلوم متوسط ( )
2- بكالوريوس ( )
3- دبلوم عالي ( )
4- ماجستير ( )
5- دكتوراه

- نوع المدرسة:

1- تعليم أساسي: 1-4 ( ) 5-10 ( )
2- تعليم عام ( )
3- تعليم عام + تعليم أساسي ( )

- جنس طلاب المدرسة:

1- ذكور ( )
2- إناث ( )
3- مختلط ( )

ثانياً - لكل فقرة مما يأتي اختر الرقم الذي يمثل مستوى تقديرك للجانب المحدد من جوانب المشروع المقوم وضع اشارة (√) بجانب ذلك الرقم .

(1) أهداف المشروع:

- مدى واقعية أهداف المشروع في تطوير العملية التعليمية التعلمية .

(5)	(4)	(3)	(2)	(1)

(2) أدوات جمع البيانات:

- ترتبط أدوات جمع البيانات بالمعايير الخاصة بكل عنصر من عناصر التقويم .

(5)	(4)	(3)	(2)	(1)

- تتطلب عملية تعبئة أدوات جمع البيانات فترة زمنية طويلة .

(5)	(4)	(3)	(2)	(1)

(3) البرامج التدريبية:

- يمتلك المدربون في برامج المشروع التدريبية الكفايات التدريبية اللازمة لعملية التقويم .

(5)	(4)	(3)	(2)	(1)

- أكسبتني برامج التدريب المهارات اللازمة لتطبيق أدوات المشروع .

(5)	(4)	(3)	(2)	(1)

(4) خطة تطبيق المشروع:

- عدد المدارس التي يتم اختيارها سنويا لتطبيق المشروع كافية.

(5)	(4)	(3)	(2)	(1)

- الفترة الزمنية (عام دراسي واحد) بين تطبيق التقويم الذاتي وتطبيق التقويم الخارجي للمدرسة مناسبة.

(5)	(4)	(3)	(2)	(1)

(5) التكلفة المالية:

- تفي الصلاحيات المالية الممنوحة للمدرسة بمتطلبات ومستلزمات تنفيذ المشروع.

(5)	(4)	(3)	(2)	(1)

- للمدرسة القدرة على تحمل الأعباء المالية التي يتطلبها المشروع.

(5)	(4)	(3)	(2)	(1)

(6) الإشراف والمتابعة:

- يلقى المشروع اهتماما ومتابعة مستمرة من قبل المسؤولين في الوزارة .

(5)	(4)	(3)	(2)	(1)

- يلقى المشروع اهتماما ومتابعة مستمرة من قبل المسؤولين في المنطقة التعليمية.

(5)	(4)	(3)	(2)	(1)

(7) **التغطية الإعلامية:**

- يحظى المشروع بتغطية إعلامية كافية ومتنوعة (صحافة، إذاعة، تلفاز، ندوات، محاضرات، مطويات. . . الخ).

(5)	(4)	(3)	(2)	(1)

- تعمل المدرسة على نشر ثقافة المشروع بين أفراد المجتمع المحلي.

(5)	(4)	(3)	(2)	(1)

(8) **إعداد وكتابة التقرير الختامي :**

- يشتمل التقرير الختامي على جميع عناصر التقويم (التعليم، التعلم، الإدارة المدرسية) .

(5)	(4)	(3)	(2)	(1)

- يتضمن التقرير الختامي تشخيصا واضحا لنقاط القوة والضعف ومقترحات التطوير.

(5)	(4)	(3)	(2)	(1)

مراجع الفصل الخامس

1- أبو زينه، فريد كامل، مروان الابراهيم، عامر قنديلجي، عبد الرحمن عدس، وخليل عليان (2005). **مناهج البحث العلمي (الكتاب الثالث) طرق البحث النوعي**. عمّان: دار المسيرة للنشر والتوزيع والطباعة.

2- الزاملي، علي عبدجاسم، حميراء السليمانية، وجيهة العاني، هاشم الشماسي، عبدالله الكيومي، محمد الكيومي، وعبدالله العبري (2008). **دراسة تقويمية لنظام تطوير الأداء المدرسي في مدارس سلطنة عمان**. بحث غير منشور، مسقط.

3- ثورندايك، روبرت وهيجن، واليزابيت (1989). **القياس والتقويم في علم النفس والتربية**. ترجمة عبد الله زيد الكيلاني وعبد الرحمن عدس. عمان: مركز الكتب الأردني.

4- عطيفة، حمدي أبو الفتوح (2002). **منهجية البحث العلمي وتطبيقاتها في الدراسات التربوية والنفسية**. القاهرة: دار النشر للجامعات.

5- علام، صلاح الدين محمود (2003). التقويم التربوي المؤسسي: أسسه ومنهجياته وتطبيقاته في تقويم المدارس. القاهره: دار الفكر العربي.

6- كاظم، علي مهدي، وكلش، صبيح (2006). تقويم برنامج التربية الفنية بجامعة السلطان قابوس من وجهة نظر الخريجين، **مجلة رسالة الخليج العربي**، السعودية، 28 (100)، 13-48 .

7- AL- Zameli: ALi. (1985). **The relationship between teachers - evaluation and students academic achievement in the sixth grade of Iraqi elementary schools**. Unpuplished Doctoral Desertation . University of Pittsburgh, USA.

8- Bodgan, R., & Tylor, S. J. (1975). **Introduction to Qualitative Research Methods** . New York: Wiley.

9- Creswell, J.W. (2002). **Educational Research : Planning, Conduct-**

ing, and Evaluation Quanlitative and Qualitative Research. New Jersey: Pearson Educational Reseaarch Inc.

10- Douglas, J.D. (1976). **Investigative Social Reseaarch**. Callifornia: Sage.

11- Gall, M.D., Gall, J.P., & Borg, W. R. (2003). **Educational Research: An Introduction**. (7eth ed). Boston: pearson Educational Inc.

12- Gay, L. R., Mills, G. E., & Airasian, P. (2006). **Educational Research: Competencies for Analysis and Applications**. New Jersey: Pearson Education, Lnc.

13- Guba, E. G., & Lincoln, Y. S. (1981). **Effective Evaluation**. California: Jossey- Bass Inc, Puhllishers.

14- Mead, M. (1928). **Coming of age in Samoa: A psychological study of primitive yoath for western civilization**. New York: Morrow. In Creswell, J.W.(2002). **Educational Research**. New Jersey: Pearson Educational Inc.

15- Royse, D., Thayer, B., Padgett, D., & Logan, T. (2001). **Program Evaluation: An Introduction. (3rd Ed.)** Belmont. CA: Brooks- cole.

الباب الثاني

القياس Measurement

الفصل السادس : مفهوم القياس

الفصل السابع : اختبارات التحصيل وأنواعها

الفصل الثامن : التقويم البديل لاختبارات الورقة والقلم

مفاهيم وتطبيقات

الفصل السادس

القيـــاس

مفهوم القياس

مستويات القياس

خصائص القياس النفسي

الاختبار التحصيلي

أنواع الاختبارات التحصيلية

مواصفات الاختبار الجيد

أولاً - الصدق

ثانياً - الثبات

ثالثاً - الموضوعية

رابعاً - الشمول

خامساً - التقنين

6

الفصل السادس

مفهوم القياس Measurement

مقدمة

يستخدم البعض مفهوم القياس ليعني التقويم أو التقويم ليعني القياس دون تمييز بين المفهومين اللذين على الرغم من أن بينهما علاقة قوية تتجسد في الهدف من استخدامهما الذي يتمثل في الحصول على أوصاف كمية أو الحكم على الظواهر أو البرامج بعد تحديد مستوى أدائها، لكنهما يختلفان في الأسلوب والتطبيق. وقد سبق أن أوردنا تعاريف متنوعة لمفهوم التقويم الذي يهدف إلى الحكم على الظاهرة المقوَّمة وفقاً لمعايير محددة. ولعل واحداً من تلك المعايير هو قياس أداء الفرد أو البرنامج لأجل توفير معلومات كمية ننطلق وفقاً لها لإصدار أحكامنا عن تلك البرامج أو أولئك الأفراد.

مفهوم القياس

القياس لغة هو التقدير، فنقول قاس الشيء بغيره، أو على غيره أي قدره على مثاله. والقياس في التربية وعلم النفس له تعريفات عديدة منها:

1- **تعريف ستيفنس**: العملية التي يتم بواسطتها التعبير عن الاشياء والحوادث باعداد حسب شروط او قواعد محددة (Stevens, 1951)

2- **تعريف كرونباخ**: عملية اعطاء ارقام لكل فرد عن كل خاصية على مقياس مؤلف من وحدات متساوية كالطول والذكاء والاستعداد . . . الخ (Cronbach, 1970.

3- **تعريف جيس** : عملية استخدام الارقام لوصف الخصائص وصفاً كمياً، او تكراراً طبقاً لمجموعة محددة من القواعد (Chase, 1978).

4- **تعريف عبد الرحمن**: عملية اضفاء الارقام على الخصائص أو الصفات، أو هو عملية تكميم الصفات والخصائص طبقاً لقواعد محددة (عبد الرحمن، 2003).

5- وفي مجال التربية يعرف بوفام (Popham, 2000) القياس التربوي بانه العملية التي يستخدم فيها المعلمون استجابات الطلبة نحو منبهات طبيعية أو مصطنعة كإستدلالات على معارفهم ومهاراتهم ومشاعرهم.

ومن خلال استعراض التعريفات السابقة يمكن الاستنتاج على ان القياس يتصف بأنه:

أ- عملية Process، بمعنى أنه عمل منظم ومخطط، وليس عملاً عفوياً او عشوائياً.

ب- تكميم النوع، أو اعطاء الخصائص ارقاماً.

ج- يتطلب قواعد rulesأو معايير norms، لمقارنة الارقام التي تم اعطاؤها للخصائص، تتصف هذه القواعد بالخصائص الأتية:

1- التمايز: أي أن لكل عدد شكل يميزه عن عدد آخر.

2- الترتيب: إن أي مجموعة من الأعداد يمكن ترتيبها ترتيباً كاملاً فيما بينها.

3- البداية المحددة: أي أن الأعداد لها بداية ثابتة هي الصفر ثم تتدرج بعدها الأرقام السالبة أو الموجبة.

كما أننا عند استخدام الأرقام لتمثيل البيانات نستخدم واحدة أو أكثر من الخصائص السابقة، فعلى سبيل المثال، قد نستخدم الأرقام كرموز لمفردات أو لفئات كأرقام السيارات أو أرقام الدور.. الخ.

وقد نستخدم الأرقام لتمثيل ترتيب مجموعة من المفردات فيما بينها حسب مقدار خاصية معينة كما يحدث في ترتيب مجموعة من الطلبة حسب الطول، العمر، الذكاء. وقد نستخدم الأرقام لتمثيل الفروق المطلقة في مقدار خاصية معينة، مثال ذلك، الفرق بين طول علي (160سم) وطول أحمد (140سم) هو (20سم). أو أن الفرق بين ذكاء طفلين ذكاء أحدهما (110درجة ذكاء) والآخر (90 درجة ذكاء) هو (20 درجة ذكاء).

وأخيراً إننا نستخدم الأرقام لتمثيل المقادير الحقيقية المطلوبة للخاصية. ويتضح أن هذه الاستخدامات تعكس مستويات متدرجة من الدقة في تمثيل خصائص الأشياء بالأرقام.

مثال:

فلو أردنا قياس ارتفاع معين فإننا نطبق الأرقام لقواعد تتعلق بوحدات رقمية محددة لعل أكثرها شيوعاً (السنتيمترات)، حيث نطبق هذه الوحدات على جميع الأشياء التي نريد قياس ارتفاعاتها، وعندما نريد قياس السرعة أو المسافة أو الوقت فكل منها يُكمم وفقاً لقواعد تتعلق بتطبيق الأرقام، وفي القياس النفسي هناك مجموعة من القواعد تحكم تطبيق الأرقام لتَصف خصائص الأداء الإنساني، هذه الأرقام تأتي من خلال أدوات القياس التي نستخدمها لهذا الغرض سواء منها الاستبيانات أو المقاييس أو الاختبارات.

كما أننا لو دققنا النظر في مصطلحات من مثل : الطول، الوزن، التحصيل، الذكاء، وهي عبارة عن خصائص أو صفات للأفراد أو الأشياء، ولو تصورنا عدم وجود أرقام لدينا (كما هي حالة الإنسان القديم) نجد أننا لا نستطيع إلا مقارنة أو مفاضلة الأشياء أو الأفراد ببعضها وفقاً لتلك الصفات فنقول: إن علياً أطول من محمود، وأن هذا الشيء أثقل من ذاك، وأن فلاناً أذكى من فلان وإلى غير ذلك، ولكن بعد أن توصل الإنسان إلى الأرقام بمفاهيمها ورموزها تغيرت الصورة بلا شك، فأصبح يضفي تلك الأرقام على الخصائص، فيقول بما أن علياً طوله (170سم) وأن محمود طوله (165سم) فإن علياً أطول من محمود، وبما أن وزن ما في هذا الكيس يساوي (10 كيلو غرامات)، وما في الكيس الثاني يزن (5 كيلو غرامات)، فإن وزن الكيس الأول أكبر من الكيس الثاني، وهكذا بالنسبة للذكاء والتحصيل وغيرهما.

مستويات القياس: Levels of Measurement

من خلال التعريفات التي تمَّ استعراضها في الصفحات السابقة لمفهوم القياس والتي تلتقي جميعها على أن عملية القياس وصف كمي للخصائص أو السمات أي استخدام الأرقام لوصف الأحداث والأشياء طبقاً لقواعد معينة وبأداة متفق عليها، فإن ذلك يعني أنه عند تغيير هذه القواعد، أو عند

استخدام الأرقام تحت قواعد مختلفة سيؤدي إلى الحصول على أنواع مختلفة من المقاييس، أي أن جميع المقاييس سوف لن تكون من فئة واحدة، لذلك ركزَّ العاملون في ميدان القياس على تحديد أنواع مختلفة للمقاييس ودرجة ملائمتها للعمليات الحسابية المعروفة لكل نوع من هذه الأنواع وكما سنرى من خلال استعراض تلك الأنواع المعروفة وهي بأربعة مستويات طبقاً للتصنيف الذي اقترحه ستيفنس Stevens :

أ - المقاييس الأسمية :Nominal Scales

وتعتبر من أبسط أنواع المقاييس، حيث تستخدم الأرقام فيها لتدلَّ على أشياء أو مجموعات معينة، مثال ذلك عندما نحدد رقماً لكل مجموعة من الناسً أو الفرقُ أو الصفوف أو الأشياء فنقول: المجموعة رقم (1)، والمجموعةَ رقم (2)، والمجموعة رقم (3)، أو المنطقة الأولى، والمنطقة الثانية، والمنطقة الثالثة. . ، أو الصف الأول، الصف الثاني، الصف الثالث وهكذا. . ، ومن الملاحظ هنا أن الأرقام أخذت دور العناوين للدلالة على الأشياء التي تمثلها، ولا يعني في هذه الحالة استخدام عمليات الجمع أو الطرح أو الضرب أو القسمة، فنحن لا نستطيع القول أن حاصل جمع الفريق رقم (1) والفريق رقم (2) يساوي الفريق رقم (3). كما يسمى هذا النوع من المقاييس (مقياس التصنيف) لأنه يستخدم الأرقام لتصنيف الوحدات أو الأشياء وفقاً لخاصية معينة أو أكثر، ومما تجدر الإشارة إليه أن القاعدة التي يَعتمد عليها هذا المستوى من القياس تتضمن عدم إعطاء نفس الرقم للمجموعات المختلفة، أو إعطاء نفس المجموعة أرقاماً مختلفة في آن واحد. إن هذا النوع من القياس يفيد المعلم أو الباحث في التربية وعلم النفس عندما يُريد تصنيف الطلبة حسب الجنس فيعطي للذكور رقم (1) وللاناث رقم (2) مثلاً، أو في تصنيف التخصص الدراسي للطلبة إلى (دراسة علمية) رقم (1)، و (دراسة إنسانية) رقم (2).

ب - المقاييس الرتبية أو (الترتيبية): Ordinal Scales

ترتكز هذه المقاييس على أساس ترتيب الأحداث أو الأشياء أو الوحدات بناءً على خاصية واحدة مُعينة، مثال ذلك عندما نُرتب المتسابقين في نشاط

رياضي معين حسب الأولوية إلى: الفائز الأول، الفائز الثاني، الثالث. . . وهكذا. ويُلاحظ أن هذا القياس يتأثر ببداية العد أو الترقيم، على عكس القياس الأسمي السابق الذي لا يتأثر ببداية العد. ففي هذا المقياس نلاحظ أن المتسابق الأول يأخذ الرقم (1)، والمتسابق الثاني يأخذ الرقم (2). . وهكذا، وهذا يَعني أن المتسابق الأول الذي وصلَ إلى خط النهاية قبل الآخرين يملك قدرةً على الجري أكثر من الآخرين، يليه الثاني، ثم الثالث وهكذا. . . ونستنتج من ذلك أن هذا الاستخدام يدلُّ على أن الشخص الأول يملك من السمَة أو الخاصية المُقاسة أكثر مما يملكه الآخرون، ولكنه لا يدل على مدى أو مقدارَ ما يملكه كَلٌّ منهم بشكل دقيق ومحدد بالضبط. وهناك خاصية أخرى تميز هذا النوعِ من القياس، وهي أن المسافات بين الوحدات أو الأفراد أو الأشياء التي تمَّ ترتيبها لا يُشترط أن تكون متساوية على الرغم من أن ذلك التساوي قد يبدو ظاهرياً من خلال الأرقام التي تدل على تلك الوحدات أو السمات، فعلى سبيل المثال لو طُلبَ إلينا ترتيب أربعة أشخاص وفقاً لأطوالهم المقاسة بالسنتيمترات وهي (163سم، 170سم، 165سم، 160سم) فإن ترتيب الشخص الذي طوله (170سم) سيكون الأول، والذي طوله (165) سيكون الثاني، والذي طوله (163) سيكون الثالث، أما الرابع فهو الشخص الذي طوله (160)سم. لكن ذلك لا يعني بأن الفرق بين طول الأول والثاني يساوي الفرق بين طول الثالث والرابع.

ج - المقاييس الفاصلة أو (الفئوية): Interval Scales

تعتبر هذه المقاييس أكثر دقةً وملائمةً للعمليات الحسابية من النوعين السابقين، وفيها تكون المسافاتَ بين وحدات القياس متساوية، مثال ذلك ما نفترضه من تساوي المسافات على المحرار، أو البارومتر (مقياس الضغط الجوي)، أو الدرجات الناشئة عن اختبار تحصيلي في مادة دراسية معينة. ومما تجدر الإشارة إليه هنا إن التساوي بين وحدات هذه المقاييس تمكننا من أن نُحدد ما إذا كان الشيء يساوي شيئاً آخر أو أنه أكبر أو أصغر منه. كما أن المسافات المتساوية على هذا المقياس تدل على مقادير متساوية من الخاصية المُقاسة، ولذلك فإن باستطاعتنا جمع قيم هذه المسافات أو طرحها، ولكن

الشيء المميز لهذه المقاييس عدم وجود (صفر حقيقي أو مطلق)، أي أن الصفر فيها لا يعني انعدام الخاصية المُقاسة. مثال ذلك عندما نقول أن درجة حرارة الجو هذا اليوم هي (صفر)°، فإن ذلك لا يعني عدم وجود درجة حرارة إطلاقاً بل وجودها بمستوىً معين (واطئ)، كما أن حصول الطالب على درجة صفر في امتحان مادة (الكيمياء مثلاً) لا يعني أن ذلك الطالب لا يملك معلومات إطلاقاً في هذه المادة بل أن معلوماته في المادة التي امتحن فيها (الموضوعات التي شملتها أسئلة الامتحان) كانت ضعيفة، وبذلك يمكن القول إن الصفر هنا نسبياً Relative وليس مطلقاً Absolute.

وفي هذا النوع من المقاييس يمكن استخدام عمليات الجمع والطرح والضرب دون أن تغير جوهر العلاقة بين المتغيرات أما القسمة فلا نستطيع القيام بها هنا لأنها تفترض وجود الصفر النسبي، فعلى سبيل المثال لو أن درجة الحرارة في مدينة ما كانت في يوم معين تساوي (30 درجة مئوية) وفي مدينة تقع شمالها كانتَ درجة الحرارة في نفس اليوم تساوي (15 درجة مئوية) فإن ذلك لا يعني أن درجة حرارة المدينة الأولى تساوي ضعف درجة حرارة المدينة الثانية بعد قسمة درجة الحرارة للمدينة الأولى على درجة حرارة المدينة الثانية أي (15/30 = 2) أي الضعف. وبنفس المنطق فإننا لا نستطيع القول أن الطالب أحمد الذي حصل في اختبار الذكاء على (140 درجة ذكاء) يملك ضعف ذكاء محمود الذي حصل على درجة ذكاء تساوي (70 درجة ذكاء) نظراً لعدم وجود صفر حقيقي يحدد انعدام الصفة المقاسة (انعدام الذكاء).

د - مقاييس النسبة: Ratio Scales

تتميز هذه المقاييس بأن لها وحدات متساوية مثل المقاييس الفاصلة لكن الفرق بينهما أن الصفر في هذه المقاييس صفراً حقيقياً مطلقاً Absolute يُشير إلى انعدام الصفة المقاسة تماماً، ففي قياس خاصية الطول أو الوزن او الحرارة

(*) يُعد القياس المئوي والفهرنهايتي من المقاييس الفئوية، أما مقياس كلفن Kleven فيعد من المقاييس النسبية، لأن الصفر فيه صفر مطلق يُشير إلى غياب وجود الظاهرة، وهو يعادل 273 في القياس المئوي.

حسب مقياس كلفن على سبيل المثال يعني الطول صفر أو الوزن صفر عدم وجود أي طول أو وزن. ويكثر استخدام هذه المقاييس في العلوم الصرفة الطبيعية بينما يقل (أو لا يستخدم) في العلوم السلوكية، لكن استخدامها ينحصر في العلوم السلوكية كعلم النفس والتربية في مجالات ضيقة عندما نريد قياس أطوال أو أوزان الطلبة للاستفادة من ذلك في وصفهم عندما يكونوا مشاركين في عينة بحث نقوم بإجرائه في المجال التربوي.

ويمكن توضيح طبيعة الأنواع الأربعة لمستويات القياس وحدود استخدامها وأمثلة لكلٍ منها في الشكل رقم (6-1):

الشكل (6-1) الفروق بين مستويات القياس وحدود استخدام الارقام فيها.

مستوى القياس وطبيعته	حدود استخدام الأرقام في هذا المستوى	أمثلة على مستوى القياس
"المقاييس الأسمية" ● أبسط مستويات القياس ● الهدف من إضفاء الأرقام على الأشياء لأجل تمييزها عن بعضها فقط.	● لا إستخدام للعمليات الحسابية الأربعة. ● يسمح باستخدام العدّ فقط.	● أرقام التلفونات. ● أرقام الشوارع. ● أرقام لاعبي كرة القدم.
"المقاييس الرتبية" ● ترتيب الأشياء أو الأحداث بناءً على خاصية معينة.	● لا إستخدام للعمليات الحسابية الأربعة. ● تُشير الأرقام إلى رتبة الفرد في سمة أو خاصية معينة.	● ترتيب المتسابقين في ركض (100م) حسب وصولهم إلى خط نهاية السباق. ● ترتيب مجموعة طلبه حسب أطوالهم أو درجة ذكائهم.
"المقاييس الفاصلة (الفئوية)" ● هناك مسافات بين الوحدات. ● تساوي المسافات بين الوحدات. ● يكون الصفر نسبياً. أي أنه لا يعني انعدام الصفة المقاسة.	● يمكن استخدام عمليات الجمع والطرح والضرب فقط. ● لا تستخدم عملية القسمة	● درجة حرارة الجو. ● درجة ذكاء الفرد. ● درجة تحصيل الطالب في مادة دراسية معينة.

أمثلة على مستوى القياس	حدود استخدام الأرقام في هذا المستوى	مستوى القياس وطبيعته
● قياس الطول. ● قياس الوزن. ● قياس الإنتاج الزراعي.	● يمكن استخدام جميع العمليات الأربعة في هذا المستوى من القياس (الجمع، الطرح، الضرب والقسمة).	"المقاييس النسبية" ● وجود فئات متساوية بين الوحدات. ● الصفر يعني عدم وجود الصفة المقاسة. أي أنه (مُطلق).

هرمية مستويات القياس

تتميز مستويات القياس الأربعة السالفة الذكر، بأن كل مستوى يمتلك خصائص المستويات السابقة له فضلاً عن الخاصية التي تميزه عما سبقه، وهذا يشير الى مفهوم هرمية مستويات القياس. والشكل (6-2) يوضح ذلك.

الشكل (6-2) هرمية مستويات القياس وخصائصها

مستوى القياس	خصائص القياس			
نسبي	التصنيف	الترتيب	المسافة	التناسب
فئوي	التصنيف	الترتيب	المسافة	
رتبي	التصنيف	الترتيب		
إسمي	التصنيف			

خصائص القياس النفسي Characteristics of Psychological Measurement

مما لاشك فيه أن القياس النفسي يتميز بسمات مختلفة في بعض الأوجه عن القياس المادي المستخدم في العلوم الطبيعية. كما أن دقة القياس النفسي هي بلا شك أقل منها في القياس الفيزيائي (المادي)؛ ويمكن أن يعود السبب في ذلك إلى أن الخصائص الفيزيائية تتصف بالحدود المادية الماثلة للعيان في الشيء المراد قياسه. فعندما نقول على سبيل المثال أن الماء يتكون من ذرتين من

الهيدروجين وذرة واحدة من الأوكسجين، فإن المتخصص بعلم الكيمياء يستطيع أن يحدد ذلك بدقة ويفصلهما أو يضمهما إلى بعض لتكوين قطرة ماء. ولكن هذا الأمر يصعب إثباته في تحديد عناصر الشخصية بشكل دقيق ماثل للعيان، ونفس الشيء يمكن إطلاقه على عناصر الذكاء. ولعل السبب في ذلك يعود إلى أن الذكاء والشخصية والانفعال وغيرها مفاهيم ليست مادية، بل هي مفاهيم افتراضية تدل على كيانات غير مادية، ولا نستطيع أن نراها أو نحدد جوانبها؛ ولذلك فإن قياسها سيكون بشكل غير مباشر؛ أي أننا نحددها من خلال آثارها والإحساس بها اللذان ينعكسان على ردود فعل الفرد في استجابات معينة. كما أن هذه الظواهر تتصف بعدم الثبات المطلق لأنها تعتمد في وجودها على العديد من المتغيرات غير المستقرة والتي تؤدي إلى عدم استقرار تلك الظاهرة. ونتيجة لذلك فإن القياس النفسي سوف تعوزه الدقة والضبط أحياناً ولذلك وضع العلماء نسبة من الخطأ يمكن قبولها في هذا النوع من القياس.

إن أدوات القياس التي يتم التوصل إليها في القياس النفسي لا تقيس كمية الصفة المطلوب قياسها كما يحدث في قياس وزن قطعة من النحاس مثلاً، لأننا عندما نقيس خاصية أو قدرة نفسية أو عقلية فإن الأداة لا تقيس كمية تلك القدرة التي يمتلكها الفرد، وعندما نقيس العلاقة بين أي خاصية لدى الفرد وأداءه فإننا لا نقيسها بوحدات مطلقه، إنما يتم في إطار نسبي ينشأ عن الاختلاف أو التباين الموجود فعلاً بين خصائص الأفراد وأدائهم. أي أن ما تُقيسهُ أدوات القياس النفسي لقدرات أو خصائص الفرد غير المادية يجري في إطار "نسبي وليس مطلقاً" . من ناحية أخرى علينا أن نتذكر حقيقة هامة في علمِ النفس ألا وهي حقيقة (الفروق الفردية) بين الناس والتي يمكن أن تنعكس على بناء أداة القياس نفسها وتحديد وحداتها، فالأداة التي تُبنى من أجل قياس الفروق الفردية تختلف عن الأداة التي تُبنى من أجل قياس الكمية، أي أن الأداة التي تُبنى من أجل القياس النفسي، تختلف عن الأداة التي تُبنى من أجل القياس المطلق. وفيما يلي عرض لخصائص القياس النفسي والتربوي:

(1) أنه قياس غير مباشر (Indirect)

أي أن السمة أو مقدار ما يمتلكه الفرد منها تُقدر من خلال أدائه على مؤشرات محددة ذات علاقة بتلك السمة. أي أننا في قياسنا لسمة معينة لدى الفرد نستدل عليها أو على مقدارها من خلال سلوكه الظاهر. فالذكاء لا نستطيع تحديد درجته لدى فرد معين بشكل مباشر، بل نستدل على كميته من خلال إجابات ذلك الفرد على بنود مقياس ذكاء يتصف بسمات سايكومترية مطلوبة. وتحصيل الطالب الأكاديمي يحدد من خلال إجابات ذلك الطالب على أسئلة اختبار تشمل المادة الدراسية التي سبق وأن درسها. وهكذا بالنسبة لتحديد مستوى انفعال الفرد ونوع أو قوة اتجاهاته نحو موضوعٍ معين أو درجة قلقه من الامتحانات.

(2) إنه قياس غيرتام (Incomplete)

أي أننا من خلال عملية قياس سمة معينة نقوم بقياس مجموعة جزئية من المثيرات التي تَستَثير استجابة الفرد (أي سلوكه) وهذه المجموعة لا تمثل جميع المثيرات أو كلها المتعلقة بالصفة المقاسة. فالاختبار التحصيلي الذي يتضمن مجموعة من الأسئلة لا يمثل جميع الأسئلة التي تكوّن إجاباتها سلوكاً كاملاً لما يمكن أن يحمله الطالب لتعطي صورة كاملة عن تحصيله بل تمثل إجاباتها جزءاً من كل. وبذلك يكون قياس التحصيل هنا قياساً غير كامل.

(3) إنه قياس نسبي وليس مطلق (Relative)

أي أننا لا نستطيع أن نفسّر أو نحكم على الدرجة التي يحصل عليها الطالب في مادة الجغرافيا مثلاً على أنها جيدة أو ضعيفة إلامن من خلال مقارنتها بمعيار معين كأن يكون ذلك المعيار (متوسط علامات المجموعة) التي شاركها الطالب نفس الامتحان حيث يكون الحكم على مستوى تحصيل الطالب من خلال مقارنة علامته بمتوسط علامات زملائه، فإما أن تكون أعلى من المتوسط أو أقل منه ويجري الحكم وفقاً لتلك المقارنة.

(4) قد تصاحب عملية القياس أخطاء معينة.

من الأمور الأخرى المهمة في مجال القياس النفسي والتي اهتم بها علماء

النفس أكثر من غيرهم موضوع (أخطاء القياس Error of Measurement) التي تؤدي إلى اللبس والغموض في التفريق بين درجة القياس الحقيقية ودرجة القياس الظاهرة أو المستخدمة وهو ما أدى بالعلماء إلى تحديد ثلاثة أنواع أو مصادر للخطأ تؤثر بشكل واضح على نتائج القياس مما يستدعي الاحتياط لها وهي:

أ. الخطأ الثابت: **Systematic Error**

وهو الخطأ الذي يعود إلى المقياس المستخدم نفسه ويتكرر بصورة منتظمة وله نفس التأثير على كل درجة على نفس المقياس. فلو استخدمنا مسطرة قياس فيها زيادة أو نقصان قدره (1سم)، يصبح سهلاً بعد معرفتنا به أخذه بنظر الاعتبار لتحديد القياس الحقيقي وسوف لن يشكل مشكلة كبيرة لعملية القياس.

ب. خطأ القياس: **Measurement Error**

وهو الخطأ الناتج عن استخدام الدرجة الظاهرية في القياس بدلاً من الدرجة الحقيقية، وهو نوع من الخطأ يحتاج إلى معالجة إحصائية للتحكم فيه.

جـ. الخطأ العشوائي: **Random Error**

وهذا الخطأ لا يمكن ضبطه أو السيطرة عليه لأنه عشوائياً، لكننا من خلال زيادة حجم العينة يمكن السيطرة عليه (عبد الرحمن، 2003).

(5) في القياس النفسي والتربوي لا يعني (الصفر) انعدام الخاصية المقاسة

فلو حصل الطالب على درجة صفر في امتحان مادة معينة، فلا يعني ذلك أن هذا الطالب لا يملك أية معلومات في تلك المادة، وإنما معلوماته غير كافية حول أسئلة الامتحان في تلك المادة.

وأخيراً لابد من الإشارة إلى أن القياس في أي مجال نفسي أو مادي لابد أن يتضمن ثلاث خطوات عامة هي:

1- التعرف على الصفة أو الوظيفة التي نُريد قياسها وتحديدها.
2- تحديد عدد من العمليات التي من خلالها يمكن أن تعبر الصفة عن مظاهرها لتصبح قابلة للملاحظة.
3- تحديد بعض الإجراءات والتعريفات لترجمة المشاهدات إلى صيغ كمية تعبر عن مقدار الصفة المراد قياسها .(Thorndike & Hagen, 1986)

ومما لاشك فيه أن فهم كل خطوة من الخطوات السابقة والصعوبات التي تنشأ عنها تقدم لنا فهماً سليماً لأساليب القياس النفسي ومشاكله في المجال التطبيقي.

الإختبار Test

لو عدنا إلى تعريفات القياس التي قدمت في الصفحات السابقة والتي أجمعت على أن القياس هو تحويل أو وصف الخصائص وصفاً كمياً وفقاً لقواعد محددة، لتبادر إلى أذهاننا أن الوصف الكمي للظواهر أو الخصائص يحتاج بلاشك إلى وسيلة أو أداة موضوعية متفقٌ عليها وتتضمن نظاماً رقمياً محدداً تلك الأداة هي ما يُسمى في القياس النفسي مقياساً أو اختباراً وتأخذ أشكالاً تتفق مع طبيعة الشيء المراد قياسه، فالطول مثلاً لا يمكن أن يُقاس إلا بالمتر أو السنتيمترات، والوزن لابُدَّ أن يُقاس بالكيلو غرامات، والتحصيل يقاس بالاختبار التحصيلي، وكذلك الحال بالنسبة للذكاء الذي يُقاس بما يسمى مقاييس أو اختبارات الذكاء، وهكذا يتضح لنا جلياً أن الأختبار هو أداة منظمة تستخدم لوصف الأداء أو الخاصية النفسية وصفاً كمياً.

لقد وردت في أدبيات القياس تعريفات متعددة للأختبار، فأنستازي وأوربينا (Anastasi & Urbina, 1997) تعرف الاختبار النفسي بأنه مقياس موضوعي مقنن لعينة من السلوك، شأنه شأن الاختبارات في العلوم الأخرى يتمثل في ملاحظات تركز على عينة صغيرة من سلوك الأفراد يتم اختيارها بعناية. ويعرفه جيس (Chase, 1978) بأنه طريقة منظمة لمقارنة أداء الفرد مع معيار مصمم للأداء.

وفي الوقت الذي يشير فيه التعريفان السابقان إلى أن الاختبار أداة تحكمها مواصفات معيارية موضوعية لتحديد عينة من السلوك سواء بشكل أداء أو قدرة محددة، نرى أن عوده (2002) يحدده بأداة قياسٍ منظمة تتضمن مجموعة من الإجراءات ذات قواعد وشروط محددة لتحديد درجة امتلاك الفرد لسمة أو قدرة معينة تتمثل في إجاباته على عينة من الأسئلة ?أو المواقف التي تمثل السمة المراد قياسها.

ومهما أستعرضنا من تعريفات أخرى للاختبار فلابُدَّ أن نَتوصل إلى أنه يعني أداة منظمة يتم تصميمها وفقاً لقواعد علمية موضوعية وتهدف إلى قياس سمة أو قدرة أو أداء فردي أو جمعي معين. وإذا ما علمنا أن أداء الفرد يتمثل بنشاطات أو سلوكيات متنوعة فإن واحداً منها هو التحصيل الدراسي، الذي كثيراً ما يقترن الاختبار به، كما يُسمى في أحيان كثيرة بالامتحان المدرسي، الذي تستند إليه العديد من القرارات التربوية التي تُتخذ في مدارس اليوم، حيث ينقل الطالب إلى صف دراسي أعلى أو يبقى في صفة بناء على حسن أدائه أو فشله في تحقيق الأداء الجيد في ذلك الامتحان، وأصبح هدفاً لدى البعض بدلاً من أن يكون وسيلةً، وأصبح الطالب يحلم أن يحقق أهدافه من خلال حصوله على علامات عالية في الامتحان، بل أصبح أحياناً سيفاً بيد بعض المعلمين يهددون به طلابهم كلما أظهروا نوعاً من السلوك الذي لا يرضون عليه حتى أصبح شبحاً مخيفاً ومصدراً رئيسياً لقلقهم في مختلف المراحل الدراسية. فما هو الاختبار التحصيلي؟ وما هي أهدافه؟ وما هي أنواعه؟

الأختبار التحصيلي Achievement Test

يُشير جيس (Chase, 1978) إلى أن الاختبار التحصيلي يُمثل مجموعة من الأساليب التي تُستخدم لتحديد مدى تحقق الأهداف المرسومة في العملية التعليمية، ويعرفه عوده (2002) بأنه طريقة منظمة لتحديد مستوى تحصيل الطالب في معلومات دراسية سبق وأن تعلمها بشكل رسمي من خلال إجابته على عينة من الأسئلة التي تمثل محتوى تلك المادة.

أما عدس (1991) فيعرفه بأنه أداة أو طريقة تقدم للمفحوص سلسلة من المهمات التي يُفترض أن يجيب عليها بطريقة تدل على مقدار ما عنده من تلك الصفة. ويمكن تعريف الأختبار التحصيلي بأنه وسيلة منظمة تتضمن تقديم أسئلة متنوعة الأهداف لقياس ما تعلمه الطالب في مادة دراسية درِّست له خلال فترة زَمنية معينة بهدف تحديد مستوى أدائه.

ويعتبر الاختبار التحصيلي أو ما يسمى شيوعاً بالامتحان أحد عناصر

التقويم الأساسية لأداء الطلاب في مدارسنا، وقد كان الأوحد في الساحة التقويمية قبل أن ترتفع أصوات الباحثين والعلماء عالياً لتُنَبه إلى ضرورة اعتباره واحداً من أساليب متنوعة أخرى تُستخدم لتقويم الأداء، كفاعلية ونشاط الطالب في الصف وقدرته على إنجاز المهارات التي يكلف بها وإعداده للبحوث والتقارير العلمية وغيرها من الأنشطة التعليمية الهامة في الكشف عن مستوى قدرات الطالب.

وظائف الاختبارات التحصيلية:

مهما اختلفت درجة استخدام الاختبارات التحصيلية من قبل المعلمين فإنها تؤدي وظائف عدة، سبق وأن أشرنا إلى بعضها في الحديث عن أهمية التقويم لكن هذه الاختبارات تقدم وظائف ايجابية أخرى منها:

(1) قياس التحصيل:

لابد لأي معلم أو مدير مدرسة أن يهتم بشكل خاص بتحديد مستوى أداء أو تحصيل الطلبة حتى يستطيع أن يتخذ القرارات المناسبة؛ ولذلك يلجأ المعلمون إلى استخدام أساليب متنوعة لتحديد هذا الأداء ومن بينها الأختبارات التحصيلية التي تزودهم ببيانات أو مؤشرات كمية ضرورية لإصدار أحكام أو اتخاذ قرارات معينة بشأن طلبتهم.

(2) تنشيط دافعية الطالب:

بما أن دافعية الفرد تتأثر بالعوامل الداخلية والعوامل الخارجية المحيطة به، فإن لإختبارات التحصيل المدرسية دوراً مهماً في زيادة الدافعية للطالب، سواء حقق مستوىً جيداً أو ضعيفاً في تلك الاختبارات. فلو حصل طالب ما على درجة عالية في أختبار فصلي معين، فإن ذلك يدفَعُه أكثر إلى مواصلة جهوده للمحافظةً على هذا المستوى في الاختبارات القادمة، بل وقد يطمح بعض الطلبة إلى تحقيق مستويات أعلى من ذلك. أما الطالب الذي يحقق مستوىً ضعيفاً في الاختبار فقد يراجع نفسه ويبحث عن العوامل الداخلية أو الخارجية التي أدت إلى هذا المستوى المنخفض من الأداء ووضع خططٍ كفيلة برفعه في الاختبارات القادمة.

(3) توفير التغذية الراجعة:

تُقدم الأختبارات تغذية راجعة للطالب في التعرف على أخطائه ومحاولة تجاوزها في المستقبل، وتحديد جوانب القوة لتعزيزها لاحقاً. ولا تقتصر الفائدة التي تفرزها الاختبارات على الطالب فقط، بل أنها ضرورية جداً للمعلم في تحديد أسباب إخفاق الطلبة والصعوبات التي واجهتهم مما قد يدفعه إلى مراجعة أسلوب تدريسه أو صياغته للاختبار، حيث أن المناقشات التي يُثيرها مع الطلبة بعد الانتهاء من تصحيح أوراق إجاباتهم توفر له الكثير من المعلومات التي ربما أغفلها في تدريسه ووضع المعالجات اللازمة لها بما يؤدي إلى تحسين عمله في المستقبل.

(4) تقويم المناهج الدراسية:

يمكن للاختبارات التحصيلية ولاسيما من خلال تحليل نتائجها أن توجه الأنظار إلى صعوبات قد يواجهها الطلاب بسبب المادة الدراسية المتمثلة بعدم وضوح بعض موضوعات الكتب المدرسية أو إخفاق المؤلفين في إيصال الفكرة العلمية للطالب بأسلوب يتناسب مع قدراته العقلية أو عمره الزمني، ومن خلال تلك الملاحظات والمناقشات التقويمية التي يُثيرها المعلمون يمكن إعادة النظر في بعض المواد الدراسية أو موضوعاتٍ محددة في تلك المواد.

(5) المقارنة بين الطلبة:

توفر نتائج الاختبارات التحصيلية الفرصة للمعلم أو مدير المدرسة في إجراء مقارنات بين مجموعات الطلاب وفقاً لتلك النتائج، أو لترتيبهم وتحديد المتميزين منهم أي الذين حققوا مستويات أداء عالية وتكريمهم لضمان استمرارهم على ذلك الأداء، كما تؤدي إلى تشخيص الطلاب الضعفاء ووضع الخطط اللازمة لرفع مستوياتهم.

(6) التشخيص قبل البدء بالتدريس:

يحتاج بعض المعلمين إلى التعرف على مستوى المجموعة التي يقومون بتدريسها قبل البدء بعملية التدريس لذلك يقوموا بإخضاع تلك المجموعة إلى

عملية اختبار، ومن خلال تحليل نتائجه يستطيع المعلم أن يحدد نقطة انطلاقه والأساليب الواجب اتباعها في تدريس تلك المجموعة، وبذلك يصل إلى تحقيق أهداف المادة بشكل أفضل وأكثر دقةً.

(7) القبول والانتقاء:

تقوم بعض المدارس أو الكليات بإخضاع الطالب الراغب في القبول أو الانتقال إليها لاختبارات خاصة Placement Test تهدف إلى تحديد مستوى قدراته ومعلوماته عن التخصص الذي يريد دراسته ?لاسيما في الكليات- قبل إعطاء الموافقة بالقبول فيها.

ومقابل تلك الايجابيات والفوائد التي يمكن ان نجنيها عند استخدام اختبارات التحصيل في العملية التعليمية، لا بد من التعرض الى بعض السلبيات او نقاط الضعف التي يمكن ان تبرز عند المبالغة في استخدامها، والتي ينبغي على المعلم والمدير معرفتها من اجل التقليل من دورها او منعها، ومن تلك السلبيات ما يلي:

1- ان المؤسسة التعليمية التي تتشدد وتبالغ بتطبيق الامتحانات يسود جوها نوع من التوتر والقلق العام.
2- هناك نسبة من الافراد في المجتمع هم ضحايا الامتحانات، ربما بسبب عدم عدالتها.
3- ان نظام الامتحانات بصيغته الحالية لا يهتم بجميع الطلبة، لا سيما ذوي القدرات المتوسطة وما دونها.
4- بسبب الامتحانات (أو المبالغة في استخدام الامتحانات)، أهمل المعلمون حاجات الطلبة وبالغوا في الاستعداد للامتحانات.
5- كما أن مؤسسات التعليم اصبحت تعد الطلبة لاجتياز الامتحانات اكثر مما تعدهم لعملية التعلم ذاتها.
6- للرسوب في الامتحانات آثار سلبية على نفسية الطالب، مثل انعدام الثقة بالنفس، والميل الى العزلة، والشعور بالقلق.
7- ان نظم الامتحانات الحالية (المستخدمة في المدارس)، ربما تؤدي الى تدمير احترام الذات وتقضي على التفكير لدى كثير من الطلبة.

أنواع الأختبارات التحصيلية:

لم تعد الأختبارات التحصيلية محددة بالامتحانات التحريرية فقط، بل تعددت وتنوعت أغراضها والوظائف التي تقدمها للعملية التدريسية وكذلك وفقاً لأساليب إدارتها أو تصميمها، وقد وضع المهتمون بالقياس والتقويم التربوي تصنيفات عديدة للاختبارات التحصيلية نقدم في الصفحات التالية بعضاً منها كالآتي:

1 - تصنيف الاختبارات التحصيلية على أساس تتابعي:

وفيه تصنف الاختبارات إلى ثلاثة أنواع هي:

أ. الأختبارات القبلية: **Pre tests**

ويجري تطبيق هذا النوع من الاختبارات قبل البدء بعملية التدريس بهدف الكشف عما يحمله الطلاب من متطلبات سابقة وما هي خلفياتهم المعرفية السابقة، وتستخدم نتائج هذه الاختبارات لتوجيه المعلم لأعادة النظر في الأهداف الموضوعة مسبقاً أو أسلوب التدريس وخطته، ولعل أفضل الاختبارات المناسبة لهذا الغرض هي اختبارات الاستعداد التي تهدف إلى تحديد مدى استعداد الطلاب لبدء عملية التعلم.

ب. الاختبارات البنائية: **Formative tests**

وهي الاختبارات التي يستخدمها المعلم أثناء عملية التدريس بشكل دوري كالاختبارات الشفوية أو التحريرية القصيرة (Quizes)، حيث تزود المعلم بمعلومات مستمرة وسريعة عن سير التدريس والتعلم سواء ما يتعلق بفاعلية طرق التعلم، أو تحديد نقاط الضعف والقوة عند الطلاب مما يجعلها توفر تغذية راجعة تفيد في تطوير عملية التعليم والتعَلم للمرحلة القادمة.

جـ. الاختبارات الختامية: **Summative tests**

وهي التي تجري في نهاية عملية تدريس وحدة تَعليمية أو فصل دراسي أو سنة دراسية، وتهدف إلى الحكم على العملية التدريسية بكل عناصرها بما فيها المعلم والطالب واتخاذ قرارات معينة.

2 - تصنيف الأختبارات حسب نوع فقراتها وطريقة التصحيح.

حيث تصنف إلى نوعين رئيسيين هما :-

أ. اختبارات موضوعية: Objective tests

وهي الأختبارات التي لا تحتاج من الطالب الكتابة أثناء الإجابة بل وضع إشارات أو كلمات محددة في أماكن خاصة. وقد سميت بالموضوعية لأنها لا تتأثر بالعوامل الذاتية للمصحح، وأن الطالب سيحصل على نفس العلامة اذا كان مصحح أوراق إجابات الطلبة. والاختبارات الموضوعية على عدة أنواع مثل: (أسئلة التكميل، الصواب والخطأ، الاختيار من متعدد، والمزاوجة). وسنأتي لاحقاً على توضيح تلك الأنواع مَعَ الأمثلة ومميزات واستخدامات كل نوعْ منها.

ب. اختبارات المقال: Essay test

وهي الاختبارات التي يقوم فيها الطالب كتابة إجاباته بما يشبه المقال. وتتأثر هذه الاختبارات بالعوامل الذاتية للمعلم لذلك لا تتصف علامة الطالب بالثبات إذ أنها تتغير من مصحح إلى آخر. كما أن من سمات هذه الاختبارات صعوبة تشابه إجابات الطلبة بسبب غموض أسئلتها وسعتها بحيث تختلف إجابات كل طالب عن إجابات غيره على نفس السؤال.

3 - تضيف الاختبارات وفقاً لتفسير نتائجها

وتصنف الاختبارات إلى نوعين هما:

أ. اختبارات معيارية المرجع: Norm reference test

وفي هذا النوع يتم تفسير علامة الطالب من خلال مقارنتها بعلامات زملائه في المجموعة التي خضعت لنفس الاختبار، وتسمى تلك المجموعة بالمجموعة المعيارية norm group فمثلاً إذا أردنا أن نفسر أو أن نحكم على أداء الطالب الذي حصل على علامة مقدارها (75) في مادة الرياضيات، فعلينا تحديد موقع علامة هذا الطالب بالنسبة للمتوسط الحسابي لعلامات مجموعته وبذلك يكون المتوسط هو المعيار Norm حيث نقول أن علامة الطالب أعلى أو أقل من متوسط علامات المجموعة التي ينتمي إليها.

ب . اختبارات محكية المرجع : **Criterion reference test**

وفي هذا النوع يقارن أداء الطالب بمستوىً مُعين يتم تحديده مسبقاً دون النظر إلى أداء المجموعة التي ينتمي إليها، مثال ذلك، إن الطالب الذي يحصل في امتحانٍ ما على علامة (60) يعتبر ناجحاً، أو أن الطالب الذي يحقق معدلاً في جميع المواد الدراسية للمدرسة الثانوية مقداره (75) يقبل في كلية التربية في جامعة معينة، وهنا نلاحظ أن مستوى أداء الطالب أو مستوى كفايته يتوقف على مقارنته بمستويات أداء خارجية (محكات) تمثل الحد الأدنى من الإتقان اللازم تحقيقه في سلوك الطالب لكي يُعتبر ناجحاً في تلك المهارة، ومن أشهر تطبيقات التقويم المحكي المرجع هو التعلم القائم على التمكن أو الإتقان Mastery Learning وتقويم البرامج Program Evaluation عندما يتم الحكم الإيجابي على البرنامج التربوي بتحقيق مستوى محدد للأهداف المعلنة لذلك البرنامج مثل (70%) منها أو أكثر.

4 - تصنيف الأختبارات وفقاً للقائم بإعدادها:

حيث تصنف إلى نوعين هما:

أ. اختبارات غير رسمية :**Informal**

وهي التي يقوم بإعدادها معلم الصف بنفسه، وتتصف هذه الاختبارات بأن معظمها قد لا يحقق الشروط المفروض توافرها في الاختبار الجيد بسبب قلة خبرة العديد من المعلمين للأساليب العلمية الضرورية لصياغة الاختبارات التحصيلية.

ب . اختبارات مقننة :**Standardized**

وهي التي يعدّها فريق متخصص في أساليب القياس والتقويم ويجري تجريبها على عينات كبيرة من الطلبة ويتم تحليل نتائجها وفقاً للأساليب الأحصائية المطلوبة، ويبقى تداولها واستخدامها محصوراً بين الباحثين والمعلمين الذين يسمح لهم بشكل رسمي في اقتنائها وتطبيقها وفق شروط محدودة.

5 - تصنيف الاختبارات حسب الوقت المحدد للإجابة

وهي على نوعين هما:

أ. اختبارات السرعة: Speed test

وهي الاختبارات التي تتطلب من الطالب الإجابة على أكبر عدد ممكن من أسئلتها ضمن فترة زمنية محددة، وبذلك تكون السرعة هي العامل الحاسم في تحديد أداء الفرد. مثال ذلك امتحانات اللغة الإنكليزية TOEFL و IELTS

ب. اختبارات القوة: Power test

وفيها يعطى الطالب زمناً شبه مفتوح للإجابة عليها بحيث يكون قادراً على الإجابة على جميع أسئلة الأختبار، إلا أن صعوبة الأسئلة ومستوى قدرات الطالب هي التي تحدد مستوى أدائه في ذلك الأختبار.

مواصفات الأختبار الجيد:

سبق وأن تطرقنا إلى أن الأختبار أو المقياس يمثل أداة القياس التي تعتمد عليها عمليتي القياس والتقويم، وأن أي خطأ أو قصور في تلك الأداة ينعكس سلباً على كلا العمليتين، ولذلك أصبح من المهم جداً الاهتمام بصياغة أو تطوير تلك الأداة التي تتمثل في الاختبارات على وجه الخصوص، ووضع الباحثون شروطاً ومواصفات يجب توافرها في الاختبار ليكون مؤهلاً لتوفير بيانات دقيقة وموضوعية يتم الاعتماد عليها في وصف الأشياء المقاسة، وفي الحكم على أداء الأفراد ولاسيما الطلبة أو اتخاذ قرارات صائبة بشأنهم.

لقد ترتب على أهمية الاختبار أو أي أداة قياس الاهتمام بجملة خصائص ومواصفات وشروط أساسية وتطبيقية منها على سبيل المثال لا الحصر سهولة التطبيق، وقلة التكاليف، وسهولة التصحيح، وسهولة التفسير وفقاً لمعايير واضحة ومحددة تمكن مستخدم ذلك الاختبار من تفسير الدرجات التي يحصل عليها الممتحنون ومقارنتها بعينة التقنين. أما الخصائص الأساسية التي لابد وأن تتوافر في الاختبار التحصيلي فهي الصدق Validity والثبات -Relia

bility وكذلك الموضوعية Objectivity والشمول Comprehensiveness، وسنقدم في الصفحات القادمة توضيحاً مفصلاً لهذه المواصفات وإجراءات التحقق منها.

الصدق Validity

على الرغم من وجود العديد من التعريفات المتعلقة بهذه الخاصية إلا أن أبسط تعريف للصدق يتمثل في أن الاختبار الصادق هو الذي يقيس ما وضع لقياسه وليس شيئاً آخر. فالاختبار الذي يُعَّد لقياس القدرة القرائية للطالب يجب أن يقيس هذه القدرة فعلاً ولا يتطرق لقياس قدرة الطالب على التعبير مثلاً. ويلاحظ أن هذا التعريف ينحصر في جانب المحتوى بينما يطرح باحثون آخرون تعريفات أخرى تتجاوز المحتوى إلى أمور أخرى يرونها ضرورية لتوضيح مفهوم الصدق. فجيس (Chase, 1978) مثلاً يعرف الصدق بقوله أن الاختبار الصادق هو الذي يوفر لنا معلومات تفيدنا في وضع القرارات المتعلقة بالسلوك الصفي. فمثلاً يكون اختبار الذكاء صادقاً عندما يوفر لنا درجات تساعدنا في أن نقرر كيف يؤدي الفرد المهمات التي نعتقد أنها تتطلب قدرة فكرية معينة. ويكون الاختبار التحصيلي صادقاً عندما يوفر لنا درجات تساعدنا في أن نقرر كيف أن الطالب يفهم موضوع الدرس بشكل متقن. وهناك من يربط موضوع الصدق بقدرة الاختبار على التمييز بين طرفي القدرة أو المهارة التي يقيسها، أي قدرته على التمييز بين الأداء الأعلى والأداء الأقل وكيف أنهما يختلفان عن الأداء المتوسط بقدر ما. فإذا كانت نتائج قياساته لتلك المستويات من الأداء متقاربة دل ذلك على ضعفه على قياس وتحديد الأداء بشكله الصحيح أي أنه لم يقم بالوظيفة الأساسية في عملية القياس، وبذلك سوف لن يكون اختباراً صادقاً. ولتوضيح ذلك لابد أن نستنتج أن الميزان الذي إذا وضعنا عليه قطعتين من الحجر إحداهما كبيرة والثانية صغيرة وأعطى نفس المستوى من الوزن (1/2 كغم) مثلاً فإن ذلك الميزان لا يمكن الوثوق به لأنه لم يميز بين الفروق في الأوزان. وكذلك نفس الحالة تنطبق

على مقياس ضغط الدم الذي يعطي نفس الدرجة لشخص مريض بارتفاع ضغط الدم وعمره في الخمسينات، وشخص آخر لا يعاني من هذا المرض وهو في العشرينات من عمره.

أنواع الصدق (دلالات إيجاد الصدق):

تُطرح في الأدبيات مصطلحات متنوعة تتعلق بأنواع الصدق منها صدق التمييز، والصدق العاملي، والصدق المنطقي، والصدق التجريبي، والصدق الذاتي، وإلى غير ذلك وجميعها تشير إلى نوع العملية أو الإجراء الذي يُتبع للحصول على مؤشرات عن صدق الاختبار. ولإزالة التشويش الذي يكتنف مفهوم الصدق في ذهن القارئ لاسيما الطالب المبتدئ في مجال القياس والتقويم نود توجيه الانتباه إلى أن رابطة علماء النفس الأمريكيين (American Psychological Association) حددت ثلاثة مصطلحات رئيسية للصدق هي: صدق المحتوى Content Validity، والصدق المرتبط بمحك Criterion related Validity، وصدق البناء Construct Validity، لكننا نرى أن من الضروري إضافة أنواع أخرى للصدق مثل الصدق الظاهري والصدق العاملي وإن كانا يرتبطان نظرياً بصدق المحتوى والصدق البنائي.

1 - الصدق الظاهري: Face Validity

يعتبر الصدق الظاهري من أبسط أنواع صدق الاختبار نظراً لبساطة الإجراءات التي تستخدم لتحقيقه، والتي تتطلب عرض الاختبار بصيغته الأولية على مجموعة من الخبراء أو المتخصصين من ذوي العلاقة بموضوع الاختبار وحتى المدرسين بالنسبة إلى الاختبارات التحصيلية المدرسية ولذلك يسمى أيضاً بصدق الخبراء أو صدق المحكمين. وبناءً على اتفاق أولئك الخبراء يمكن التوصل إلى صدق الاختبار الذي يعرف وفقاً لهذا الإجراء بأن الاختبار يبدو ظاهرياً وكأنه يقيس ما وضع لقياسه. ونظراً لأن هذا النوع يعتمد على رأي الخبراء فإنه سيكون عرضةً للعوامل الذاتية لهؤلاء الخبراء، لذلك فإن من

الضروري للباحث أن يوفر قدراً من المعايير الكمية البسيطة التي تسمح له بتقليل أثر تلك العوامل، وهنا يمكن للباحث استخدام أحد الأساليب الآتية لإيجاد الصدق الظاهري:

1- يمكن الاعتماد على نسبة إتفاق بين المحكيين لتكون معياراً لقبول أو حذف الفقرة التي يتفق عليها الخبراء، كأن يستخدم نسبة إتفاق لا تقل عن 70% أو 80% لهذا الغرض. فإذا كان عدد الخبراء مثلاً عشرة، وأفاد سبعة منهم أن السؤال الأول لا علاقة لهُ بموضوع الاختبار، فعلى الباحث أن يحذف ذلك السؤال، وهكذا.

2 - أما الأسلوب الكمي الثاني الذي يمكن للباحث الاعتماد عليه لإيجاد الصدق الظاهري للمقياس أو الاختبار فهو (الوسيط) من خلال الخطوات التالية:

أ. بعد أن يعد الباحث الصوره الأولى للمقياس أو الاختبار يعرضه على عدد من الخبراء المتخصصين في الموضوع والمتخصصين بأساليب التقويم والقياس وبعدد يفضل أن يكون بحدود (عشرة أو أكثر)، ويطلب من كل خبيرٍ أو (محكِمْ) على حدة أن يحدد رأيه عن مدى علاقة السؤال بالموضوع المقاس على سلم تقدير يتراوح بين (صفر إلى عشرة).

ب. يقوم المحكم بقراءة السؤال (الفقرة) ويؤشر على سلم التقدير بإشارة (√) تحت الرقم الذي يمثل رأيه بملائمة السؤال للموضوع.

جـ. بعد أن ينتهي من جمع آراء المحكمين وتقديراتهم حول اسئلة الاختبار تصنف تلك التقديرات بجدول خاص وفقاً لعددهم في كل مستوىً من مستويات التقدير من (صفر - 10)، ثم تحسب النسب المئوية للتقديرات.

مثال:

لو فرضنا أن الباحث عرض مقياساً لقياس ضغوط العمل لدى معلمي المدارس الإعدادية على عشرين محكماً وكانت تقديراتهم على إحدى فقرات المقياس المذكورة أدناه كما في الجدول الموضح.

مشكلات الطلاب أثناء الاستراحة بين المحاضرات توتر أعصابي.

سلم التقدير / المحكمين	0	1	2	3	4	5	6	7	8	9	10
عدد المحكمين وفقاً لمستوى التقدير	-	-	-	-	-	1	1	1	12	2	3
النسبة المئوية لتقديراتهم	0	0	0	0	0.05	0.05	0.05	0.05	0.60	0.10	0.15

تُحسب درجة صدق السؤال من خلال المعادلة التالية:

$$ص = ح + \frac{0.5 - مج\ ن}{ن}$$

حيث أن:

ص = درجة صدق السؤال (وهو ما نريد إيجاده).

ح= الحد الأدنى للفئة الوسيطية (الفئة التي يقع فيها الوسيط) ولتحديد الفئة الوسيطية نلجأ إلى إيجاد التكرار المتجمع الصاعد ثم معرفة ترتيب الوسيط بعد قسمة ذلك التكرار على (2). ويحسب التكرار المتجمع الصاعد من خلال جمع التكرارات مع بعضها؛ أي جمع تكرارات تقديرات المحكمين وتقسم على اثنين (20 ÷ 2 = 10) نفتش عن التكرار (10) لتحديد الفئة التي يقع فيها لتكون هي فئة الوسيط. ومن المثال السابق نجد أنها الفئة (8) من سلم التقدير والتي يكون فيها مجموع التكرارات من بداية السلم (12) أي أن (10) تقع ضمن هذه التكرارات

علماً بأن الحد الأدنى لهذه الفئة هو (7.5).

مجـ ن = مجموع النسب التي تقع قبل الفئة الوسيطية أي:

(صفر + صفر + صفر + صفر + صفر + 1 + 1 + 1 = 3)

ن = النسبة الوسيطية (أي النسبة المئوية لعدد الحكام في فئة الوسيط ومقدارها (60%).

وعند تطبيق المعادلة السابقة نصل إلى درجة صدق الفقرة وهو:

$$ص = 7.5 + \frac{0.15 - 0.05}{0.6}$$

$$= 7.5 + 0.58$$

$$= 8.08$$

وهكذا يتم حساب درجات الصدق لجميع اسئلة الاختبار.

هـ. بعد إيجاد درجات صدق الاسئلة يتم ترتيبها تنازليا من أعلى درجة إلى أقل درجة ونَقوم بأخذ الثلث الأعلى للأسئلة المرتبة طبقاً لدرجات صدقها لتشكل الاختبار بشكله النهائي بعد تحديد الصدق الظاهري وتحذف الأسئلة الباقية منه لكي نضمن درجات صدق بعيدة نسبياً عن العوامل الذاتية للمحكمين وتستند إلى معايير كمية واضحة.

ويلاحظ على هذا الأسلوب إن على الباحث أن يصمم عدداً كبيراً من الفقرات لأنه في نهاية التحكيم سيحذف حوالي ثلثي تلك الأسئلة مما يتطلب منه جهداً ووقتاً أكبر في الصياغة مما يؤدي إلى عدم استخدام هذا الأسلوب من قبل بعض الباحثين.

2 - صدق المحتوى Content Validity

ويسمى (صدق المضمون) ويعني أن محتوى الاختبار بجميع فقراته يمثل السلوك الذي يقيسه ذلك الاختبار بكل جوانبه، ولذلك يهتم الباحث في هذا

النوع من الصدق بالتأكد من وجود علاقة قوية بين فقرات الاختبار ومكونات السلوك المطلوب قياسه. ففي الاختبارات التحصيلية يمكن تحقيق ذلك من خلال مقارنة أسئلة الاختبار بالمادة التعليمية التي يتضمنها الكتاب المدرسي، أو من خلال تنظيم ما يسمى بـ (جدول المواصفات) الذي سنأتي على توضيحه لاحقاً والذي يتضمن تحديداً واضحاً للمادة الدراسية من جهة، والأهداف السلوكية المطلوب قياسها من خلال الاختبار من جهة أخرى، إلا أن تحقيق هذا الهدف يكون أكثر صعوبةً في إيجاد محتوى الاختبارات أو المقاييس النفسية مثل (مقاييس الاستعدادات ومقاييس الشخصية) حيث هناك تعريفات عديدة ومتباينة أحياناً لهذه المفاهيم ولا توجد جوانب محددة لها (قائمة محتوى) كما هو الحال في المادة الدراسية المحددة المتمثلة بمحتوى المنهج أو الكتاب المدرسي.

3 - الصدق المرتبط بمحك Criterion-related validity

في هذا النوع من الصدق يتم التأكد من صدق الاختبار أو المقياس من خلال تقديم الشواهد على علاقة الاختبار المطور باختبار آخر معروف ومتصف بصدق وثبات عالٍ ويتمتع بخصائص الاختبار الجيد أو ما يسمى بالخصائص السيكومترية، وفي هذه الحالة يسمى ذلك الاختبار (المحك) وهو الذي نقارن به الاختبار الذي نقوم بتطويره. فلو قام باحث ما بتطوير مقياس لقياس الذكاء لدى طلاب المدرسة الابتدائية، فعليه أن يختار مقياساً معروفاً يتمتع بدرجة مقبولة من الصدق ويقيس نفس الخاصية لأطفال نفس المرحلة الدراسية والعمرية، كأن يختار مقياس بينيه، أو مقياس وكسلر للذكاء، ثم يطبق المقياس الذي طوره والمقياس المعروف على نفس المجموعة من الأطفال وتحت نفس الظروف النفسية والفيزيقية ويحدد علاماتهم على كلٍ من المقياسين، أي أن تكون لكل فرد في المجموعة علامتان وهما (س) و (ص)، وبعد ذلك يمكن استخدام الوسيلة الإحصائية المناسبة لتحديد مستوى العلاقة بين مجموعتي العلامات فإن كان مستوى العلاقة مرتفعاً ذو دلالة إحصائية

مقبولة دلَّ ذلك على صدق المقياس المطور، انطلاقاً من المبدأ القائل أن المقياسين متقاربين أو يقيسان نفس الخاصة. وبما أن المقياس الثاني معروف بصدقه، فإن المقياس المطور هو بالضرورة يتمتع بالصدق لأنه ارتبط بمحك (مقياس) صادق، ويسمى هذا النوع بالصدق التلازمي Concurrent Validity تمييزاً له عن نوع آخر هو الصدق التنبؤي Predictive Validity الذي سنوضحه لاحقاً.

مثال:

لنفرض أن باحثاً ما أراد أن يتأكد من صدق اختبار طوره لقياس القدرة اللغوية لدى طلبة صف دراسي معين، فاختار مقياساً معروفاً يقيس نفس القدرة اللغوية لدى طلبة نفس الصف الدراسي وهو في هذه الحالة سيكون محكاً. طبق الباحث الاختبارين معاً في نفس الوقت على مجموعة مكونة من ثمانية طلاب فحصلوا على العلامات الآتية:

اسم الطالب	علامة الطالب على الاختبار	علامة الطالب على المحك
أحمد	37	88
محمد	34	86
سيف	32	77
ناصر	26	72
سعود	22	65
راشد	21	62
محمود	19	59
مهند	16	58

والآن دعنا نستخدم معادلة معامل ارتباط بيرسون لنحدد مستوى العلاقة بين مجموعتي العلامات على الاختبار المطور وعلى المحك، فسنجد أنها بلغت (0.986) وهذا يعني أن مجموعتي العلامات حققت ارتباطاً موجباً وعالياً، مما يدعونا للاستنتاج أن الاختبار المطور صادق لارتباطه بمحك صادق.

ويلاحظ من المثال السابق أننا طبقنا الاختبارين (الجديد أو المطور) مع الاختبار المعروف الصادق (المحك) في وقت متلازم على نفس المجموعة من الطلبة، لكن هناك أسلوب آخر لإيجاد الصدق المستند إلى محك يتعلق بإيجاد قدرة الاختبار على التنبؤ بالأداء اللاحق للمجموعة في مجال معين، وبذلك يطلق على هذا النوع بالصدق التنبؤي Predictive Validity ويتم التحقق منه بتطبيق الاختبار على مجموعة من الأفراد لتحديد أدائهم في مهارة أو مهنة وغير ذلك. ويتم الاحتفاظ بالدرجات التي حصلوا عليها من ذلك الاختبار. مثال ذلك أننا نريد أن نحدد قدرة صدق الاختبار على التنبؤ بنجاح طلاب كلية التربية في مهنة التعليم في المستقبل لذلك نلجأ إلى تطبيق هذا الاختبار على طلبة الصفوف المنتهية من الكلية، وبعد أن يتم تعيين هذه المجموعة في المدارس نتابع أدائهم بعد مدة مناسبة من العمل لنحصل على تقديرات المدراء أو تقديرات المشرفين التربويين لأدائهم ثم نعمل على إيجاد معامل الارتباط بين علاماتهم على الاختبار وعلاماتهم في المهنة (التقديرات) فإذا وجدنا علاقة ارتباطية موجبة وعالية دلَّ ذلك على قدرة هذا الاختبار بالتنبؤ بأداء المعلمين في المستقبل وبذلك يمكن استخدام مثل هذا الاختبار لاختيار الطلبة المتقدمين للقبول في كلية التربية إذا كان أداؤهم مرتفعاً على هذا الاختبار، أو يمكن اعتباره أحد أساليب الاختيار والقبول في الكلية مع مؤشرات أخرى تستخدم لهذا الغرض.

وبما أن معامل الارتباط بين علامات المجموعة على الاختبار المطوَّر وتقديرات المشرفين لأدائهم لاحقاً يعتبر مؤشراً للصدق التنبؤي للاختبار المطور إلا أن استخدامه لوحده لا يكفي للتنبؤ بالقدرة التي يقيسها. لذلك يُنصح باستخدام معادلة الانحدار للتنبؤ بدرجات المجموعة اللاحقة (ص) من خلال معرفتنا بدرجاتهم على الاختبار (س).

4 - صدق البناء (التكوين) Construct Validity

يعتبر الاختبار صادقاً من حيث البناء إذا ما تحققت علاقة الاختبار بمعلومات

تمثل نظرية معينة. وتعرف النظرية وفقاً لـ (Kubiszyn and Borch, 1987) بأنها تفسير منطقي أو عقلاني يوضح العلاقات بين مجموعة من المتغيرات.

ولكي يكون الاختبار صادقاً بنائيا يجب أن يبرهن الباحث إلى أي حد يتمكن ذلك الاختبار من قياس البناء النظري للسمة المقاسة كالذكاء والدافعية والقدرة الميكانيكية وغير ذلك. إن كل بناء أو مفهوم من تلك المفاهيم يجب أن يصاغ بشكل فقرات أو أسئلة يتضمنها الاختبار بحيث تسهل عملية رصد السلوك الملاحظ للمفحوص بشكل منظم. وتُشير أنستازي Anastasi إلى أن صدق البناء يتطلب تجميعاً متدرجاً للمعلومات المتعلقة بالمفهوم المراد قياسه من مصادر متنوعة بحيث تلقي الضوء على طبيعة المفهوم وتوضحه بشكل يُتيح للباحث تحويله إلى مفردات قابلة للقياس. ومن الجدير بالذكر أن الباحث عندما يركز عمله على صدق بناء اختباره فأنه لا يزال في المرحلة التمهيدية - أي المرحلة النظرية لتطوير الاختبار أو المقياس لأنه لازال يركز عمله في التعريفات المتعددة للمفهوم المراد قياسه، وأن وجود أكثر من تعريف يُشير إلى أن العلماء ينظرون إلى المفهوم نظرات مختلفة أحياناً انطلاقاً من اختلافات في الرؤى التي تعكسها المدارس والأطر النظرية التي ينتمي إليها كل منهم، يضاف إلى ذلك أن هناك مقاييس متنوعة لقياس نفس المفهوم مما يتطلب من الباحث الاهتمام بالتعرف على أكثر من تعريف والاطلاع على أكثر من مقياس ليحدد موقعه بين تلك التعريفات المختلفة التي تمثل مدارس فكرية مختلفة.

أما الإجراءات العملية التي يُفترض بالباحث الالتزام بها لإثبات صدق بناء مقياسه فقد أشار لها جرونلاند (Gronlund, 2006) وهي:

1- ضرورة التعريف بالإطار النظري في السمة التي ترتبط بنتائج الاختبار.

2- وضع فرضيات واضحة حول نتائج الاختبار تستند إلى الأطار النظري.

3- التحقق المنطقي أو التجريبي من الفرضيات.

فلو كان الاختبار مثلاً يتعلق بقياس الذكاء وهو مفهوم سايكولوجي له أكثر من تعريف، وهناك أكثر من مقياس طوِّر لقياسه، وهناك معلومات نظرية

تصلح كمؤشرات تعدد طبيعته، وتُشير العديد من نظريات النمو إلى إنه ذو طبيعة تطورية وهو يرتبط بمتغيرات عديدة منها ما تؤثر فيه ومنها ما تتأثر به فهنا يُفترض بالباحث أن يجمع أكبر عدد ممكن من تلك المؤشرات لكي يأخذها في الاعتبار عند إثبات صدق الاختبار الذي يقوم بتطويره، كما تفيد تلك المؤشرات الباحث في صياغة فرضياته التي يمكن اشتقاقها من الإطار النظري ومنها:

- إن علامات الأفراد على مقياس الذكاء تزداد بازدياد أعمارهم.
- إن علامات الأفراد على مقياس الذكاء تتنبأ بتحصيلهم الأكاديمي.
- إن علامات الأفراد على مقياس الذكاء ترتبط ارتباطاً طردياً مع علاماتهم في اختبارات أخرى تقيس الذكاء أيضاً.
- إن علامات الأفراد على مقياس الذكاء تميز بين الموهوبين والمتخلفين عقلياً.

ويُشير (عوده، 2002) إلى أن انسجام نتائج اختبار هذه الفرضيات يوفر دليلاً مقبولاً على صدق البناء. وأخيراً لابد من الإشارة إلى أنه رغم أهمية الشواهد المذكورة سابقاً فإن (Gronlund, 2006) يؤكد أيضاً على أن هناك شواهد من أنواعٍ أخرى لابد من توفيرها للدلالة على صدق البناء منها ضرورة إجراء دراسة شاملة عن كيفية تطوير الاختبار، وكيف يعمل الاختبار في المواقف المختلفة، وهل ترتبط درجاته ارتباطاً دالاً بدرجات مقاييس أخرى. أما الأساليب الإجرائية التي يمكن من خلالها التحقق من صدق البناء لمقياس ما فيمكن حصرها بالآتي:

أ. تغيرات النمو **Developmental Change**

إن المحك الرئيسي الذي أُستُخدِمَ للتأكد من صدق عدد من اختبارات الذكاء هو التمايز في العمر، ففي اختبار بينيه للذكاء ومعظم الاختبارات الخاصة بمرحلة ما قبل المدرسة تم اختبار صدقها بهذا المحك لتقرير ما إذا كانت العلامات التي يحصل عليها المفحوصون تُبين نوعاً من الزيادة المتصاعدة مع تقدم العمر. ونظراً لأن مثل هذه القدرات يتوقع لها أن تزداد مع تقدم العمر لاسيما في مرحلة الطفولة وحتى سن الرشد، فإن نتائج الاختبار

يُفترض أن تتسق مع هذه القاعدة لكي تؤكد صدق ذلك الاختبار من حيث البناء. ومع أن أكثر الباحثين يُشيرون إلى إمكانية الأخذ بهذا الأسلوب للبرهنة على صدق بناء الاختبار إلا أن أنستازي (Anastasi, & Urbina 1997) تُشير إلى أن هذا الأسلوب قد لا يتحقق أو ينطبق على جميع الاختبارات والمقاييس حيث أن مقاييس الشخصية مثلاً أظهرت نتائج محددة وقد تكون غير دالة من خلال استخدام هذا الأسلوب، وتُضيف إلى ذلك أن عدم الوصول إلى نتائج إيجابية في هذه المقاييس أو في بعض مقاييس الذكاء قد يرجع إلى أخطاء في العينة لاسيما في إسلوب اختيارها، وترى أن إثبات صدق البناء في الاختبارات السيكولوجية والسلوكية التي تزداد بتزايد العمر يجب أن تأخذ بالاعتبار أن تلك الزيادات تتأثر بالعوامل البيئية المحيطة بالفرد والتي يُقنن فيها الاختبار مما يتطلب من الباحث الاهتمام بموضوع اختلاف البيئات والثقافات التي قد تُحفِّز وتُعَزِّز نمو الخصائص السلوكية المختلفة، ولذلك يجب أن لا نفترض أن هذا المحك (التمايز في العمر) هو محك عالمي، بل أنه شأن المحكات الأخرى يتأثر بالثقافات التي تسود المجتمعات وهي متباينة بالتأكيد.

ب. الاتساق الداخلي Internal Consistency

بما أن السمة السيكولوجية trait تمثل مجموعة سلوكات تحدث معاً نتيجةً لارتباطها ببعضها، وبما أن هذه السلوكات هي التي يجب أن تعرض في المقياس المطور بشكل فقرات أو عبارات لفظية لتشكل البناء الكلي للخاصية المقاسة، فإنه يجب أن تشكِّل لبنةً ترص مع اللبنات الأخرى لتشكل البناء المطلوب الذي يمثل المفهوم الكلي المتماسك للخاصية المقاسة. وانطلاقاً مما تقدم فإن كل لبنة (فقرة) من لبنات المقياس يجب أن تنسجم مع اللبنات الأخرى أي مع الفقرات الأخرى في تحقيق الهدف العام الذي أُختيرت لتحقيقه أي قياس الخاصية المطلوب قياسها بالاختبار كله. وأن عدم انسجامها في تحقيق هذا الغرض يعني ضرورة حذفها واستبدالها بأخرى تكون أكثر انسجاماً مع فقرات الاختبار. ولتحقيق ذلك فإن التأكد من العلاقة الارتباطية بين أداء الأفراد على هذه الفقرة وأدائهم على عموم المقياس هو المحك الذي يؤدي إلى صدق بناء ذلك المقياس.

وللتأكد من الاتساق الداخلي لفقرات المقياس يلجأ الباحثون إلى تطبيق إحدى معاملات الارتباط (وحسب طبيعة المتغيرات) للحكم على بقاء أو حذف الفقرة. ويمكن توضيح ذلك بالمثال التالي:

مثال:

لنفترض أننا نريد إيجاد العلاقة بين علامات الطلبة على السؤال (1) وأدائهم على عموم اختبار تحصيلي في مادة دراسية معينة. علماً بأن الاختبار مكون من (10) اسئلة تتراوح علامة الطالب على هذا السؤال من (1-3)، مما يعني افتراضاً أن أعلى علامة كلية يمكن أن يحصل عليها الطالب هي (30) وأقل علامة هي (10). ويمكن أن نضع العلامات التي حصل عليها كل طالب على السؤال (1) من جهة وعلامته الكلية على عموم الاختبار من جهة أخرى كما في الجدول الآتي:-

تسلسل الطالب	درجة الطالب على السؤال (١)	درجة الطالب الكلية في الاختبار
1	3	28
2	2	25
3	1	19
4	3	26
5	2	16
6	3	20
7	2	26
8	1	15
9	2	‘18
10	1	19
11	2	17
12	3	27

ولإيجاد معامل الارتباط بين العلامات على الفقرة والعلامات الكلية يمكن استخدام معامل (ارتباط بيرسون) وذلك لأن المتغيرين مستمرين. ومن المثال أعلاه نجد أن قيمة معامل الارتباط تساوي (0.648) وتمثل علاقة موجبة ومقبولة مما يسمح لنا بإبقاء الفقرة وعدم حذفها. ونستمر بإيجاد علاقات درجات الطلاب على الفقرات الأخرى مع درجاتهم الكلية على الاختبار لأجل أن نقرر بقاء أو حذف كل فقرة. وتجدر الإشارة إلى إننا استخدمنا في المثال السابق معامل ارتباط بيرسون للسبب المذكور أعلاه وهو أن علامات الطلاب على الفقرة تمثل متغيراً مستمراً (1، 2، 3). أما إذا كانت الإجابة على فقرات الاختبار تمثل متغيراً ثنائياً متقطعاً مثل (1، صفر) فإننا في هذه الحالة سنستخدم (معامل ارتباط بايسيريال) الذي سنأتي على كيفية استخدامه عندما نتطرق إلى معاملات الارتباط وكيفية تطبيقها في صفحات لاحقة من هذا الكتاب، ونعرض فيما يأتي جدولاً يمثل علامات مجموعة من الطلاب على فقرة اختبارية افتراضية وعلاماتهم الكلية على ذلك الاختبار والمكون من (30) فقرة.

تسلسل الطالب	علامة الطالب على الفقرة (2)	درجة الطالب الكلية في الاختبار
1	0	20
2	1	28
3	1	27
4	1	26
5	1	29
6	0	18
7	1	25
8	1	26
9	0	19
10	1	20

ومن المثال أعلاه يتضح أن قيمة معامل الارتباط بين علامات الطلبة على الفقرة وعلاماتهم على عموم الاختبار تساوي (0.708) وتعتبر هذه القيمة جيدة مما يعني أن الفقرة ترتبط مع العلامة الكلية ارتباطاً قوياً.

وأخيراً لابد من الإشارة إلى أن هناك العديد من الأساليب التي يستخدمها الباحثون لإيجاد الاتساق الداخلي للاختبار عدا الأسلوب الموضح سلفاً، ومن تلك الأساليب ما يأتي:

- يميل بعض الباحثين إلى طرح العلامة التي يحققها الطالب على الفقرة من العلامة الكلية على عموم الاختبار ثم نوجد العلاقة بين العلامتين.

مثال:

رقم الفقرة (1) أعلى درجة كلية يحصل عليها الطالب (20) وهو مجموع فقرات الاختبار.

علامة الطالب على الفقرة	علامة الطالب في الاختبار	علامة الطالب النهائية عدا علامة الفقرة
1	18	17 = 1-18

وهنا ستكون علامة الطالب على الفقرة هي (1) وعلامته النهائية على الاختبار بعد طرح علامة الطالب منها هي (17). ونستمر بنفس الأسلوب بالنسبة لعلامات الطلاب الآخرين ولكل فقرة من فقرات الاختبار على حدة.

- وهناك أسلوب آخر يهدف إلى إيجاد العلاقة بين المجالات أو الأبعاد الرئيسية التي يتكون منها المقياس، أي إيجاد العلاقة بين مجموع العلامات التي يحصل عليها كل طالب على فقرات البعد Dimenssion أو ما يسمى أحياناً Subtest والعلامة الكلية لأولئك الطلبة وكثيراً ما يستخدم هذا الأسلوب في اختبارات الذكاء التي تستخدم اختبارات فرعية تعود لاختبار واحد لكنها تطبق واحدةً بعد الأخرى ومنها مثلاً (معاني الكلمات، إكمال

الصور، الرياضيات.. الخ)، أو مثلاً في مقياس الضغوط النفسية يقسم الاختبار إلى عدد من الأبعاد منها (الضغوط الاجتماعية، الضغوط الأسرية، ضغوط العمل.. الخ) فإذا ما ظهرت علاقة موجبة ومقبولة بين البعد والاختبار الكلي يتم الإبقاء على فقرات ذلك البعد، وإن ظهرت العلاقة ضعيفة يتم حذف فقرات ذلك البُعد.

مثال:

ولتوضيح هذا الأسلوب نطرح المثال التالي الذي يتضمن كيفية إيجاد العلاقة بين علامة الطالب على أحد أبعاد الاختبار والعلامة الكلية على الاختبار، ولنفرض أن أعلى مجموع للعلامات على البعد يساوي (10) وأعلى علامة كلية على الاختبار هي (100).

علامة الطالب على البعد	الدرجة الكلية للطالب على الاختبار
8	65

ومن خلال إدراج علامات جميع الطلاب على البعد المحدد وعلاماتهم الكلية في الاختبار يتم اختيار معامل الارتباط المناسب وتطبيقه لتحديد قيمة معامل الارتباط بين مجموعتي العلامات لنقرر بقاء ذلك البعد بفقراته أم حذفه من الاختبار.

ومن المهم هنا أن نذكر أن معاملات الاتساق الداخلي للاختبار وباستخدام أيٍ من الأساليب السابقة الذكر تعني بالأساس مقاييس للتجانس، ولأنها تساعد في وصف البعد السلوكي أو السمة المتمثلة في الاختبار، فإن درجة تجانس الاختبار لها بعض الصلة بصدقه البنائي، ومع ذلك فإن مساهمة بيانات الاتساق الداخلي في تحديد صدق الاختبار تبقى محدودة في غياب البيانات الخارجية عن الاختبار ولا توسع معرفتنا عما يُقيسه ذلك الاختبار (Anastasi & Urbina, 1997).

جـ. المقارنة بين المجموعات المتطرفة **Extreem groups Comparison**

مما لاشك فيه أن أية قدرة إنسانية يتوزع عليها عموم الأفراد توزيعاً إعتدالياً، ويتميز بأن غالبية الأفراد يقعون في وسط التوزيع بينما تَقع فئة بسيطة في المستوى الأعلى ومثلها في المستوى الأدنى لتلك القدرة أو السمة، ولكي يكون الاختبار صادقاً يجب أن يوضح فروقاً ذات دلالة إحصائية بين الأفراد الذين يملكون أعلى وأقل مستوى من تلك القدرة تماشياً مع الحقائق النظرية التي أوضحت في المقدمة، وإذا ما فشل في التوصل إلى دلالة تلك الفروق لا يمكن الوثوق به. فلو فرضنا أننا اشترينا جهازاً جديداً لقياس ضغط الدم لدى الإنسان وأتينا بخمسة أشخاص دلت الفحوص الطبية على أنهم مصابون بارتفاع في ضغط الدم، وخمسة أشخاص آخرين مصابون بانخفاض ضغط الدم ثم استخدمنا ذلك الجهاز وأعطانا قياسات متقاربة لضغط الدم لدى أفراد المجموعتين فإننا بلا شك سوف لن نصدق أو نثق بهذا المقياس ونعتبره مقياساً يعطي قياسات غير صادقة. وانطلاقاً من هذا المثال نريد من الاختبارات النفسية والتربوية أن تميز لنا بين المجموعتين المتطرفتين في خاصيةً معينة. فلو طبقنا مقياساً ما على مجموعة من الأفراد وأخذنا علامات الثلث الأعلى وعلامات الثلث الأدنى واختبرناً الفروق بينهما وفقاً للاختبار التائي t-test وظهرت فُروق ذات دلالة إحصائية لمصلحة المجموعة العليا بمستوى لا يزيد عن (0.05) فإن ذلك الاختبار سيكون صادقاً، وأن صدقه يتطابق مع صدق المفهوم أو البناء النظري للمفهوم المراد قياسه. أما اذا كان اتجاه الفروق في اختبار "ت" لمصلحة المجموعة الدنيا، فإن السؤال ضعيف ويجب استبعاده من الاختيار، لأن تمييزه عكسي لمصلحة المجموعة الدنيا وليس العليا.

د. التحليل العاملي **Factor Analysis**

يستخدم التحليل العاملي لتحديد صدق المفهوم أو البناء الخاص بالاختبار أو المقياس الذي يتم تطويره، حيث يهتم الباحث وفقاً لهذا الأسلوب بحساب درجة تشبع الاختبار بالعامل المراد قياسه من خلال احتساب معاملات الارتباط

بين بنود الاختبار الواحد، ثم اشتقاق العوامل المشتركة بين هذه الارتباطات، فإذا توصل إلى أن اختباره مشبع بعامل واحد فقط وأن ما دون هذا العامل بقايا يكون الاختبار صادقاً. ومفهوم الصدق هنا يشبه مفهوم الاتساق أو التجانس الداخلي حيث يصل إلى معرفة العوامل المشتركة بين الفقرات، لكنه لا يتأكد مما إذا كانت هذه العوامل هي التي قُصِدَ قياسها أم لا.

يبدأ التحليل العاملي عادةً من مصفوفة ارتباطات بين عدد من الاختبارات التي تقيس مجالاً متجانساً طبقت على عينة من الأفراد، ومن خلال تحليل هذه المصفوفة عاملياً بعدد من الفئات التصنيفية المختصرة هي العوامل التي تعبر عن التباين المشترك بينّ هذه المتغيرات. فلو ظهر ?مثلاً- أن الاختبار تَشَبَّعَ بنسبة (0.72) فهذا يعني أن الاختبار يقيس هذه القدرة (العامل الذي حصلنا عليه) بمعامل صدق عاملي قدرة (0.72) ونستطيع أن نتعرف على المكونات العاملية للاختبار بحساب تَشبعاته على العوامل المختلفة والتي حصلنا عليها من تحليل مصفوفة ارتباطية لمجال مَتجانس .(Anastasi & Urbina, 1997)

إن إجراء التحليل العاملي للاختبار أصبح إجراءً سهلاً من خلال إدخال البيانات التي نحصل عليها من تطبيقه على عينة البحث، في برامج الإحصاء المتخصصة، مثلاً برنامج SPSS أو أي برنامج آخر، وبما أن التحليل العاملي يصنف على أنه احصاء متقدم، فإنه بحاجة إلى خبرة ودراية إحصائية كافية من الباحث، ويمكن الاسترشاد بالمراجع الآتية لمعرفة طريقة تنفيذه (نجيب والرفاعي، 2006؛ George & Mallery, 2003).

الثبات: Reliability

يقصد بالاختبار الثابت ذلك الاختبار الذي يعطي تقديرات أو قياسات ثابتة إذا ما كُررَ تطبيقه على نفس المجموعة مرتين بينهما فاصل زمني. ولعل خاصية الثباتِ الآنف ذكرها تظهر بشكل أكثر دقة في القياس أو المقياس المادي، أي الذي يتعامل مع خصائص ثابته ولها كيان محدد. فعندما نُقيس طول غرفة بأداة قياس الطول وهي المسطرة عادةً ونجد أنه يبلغ خمسة أمتار ثم

نأتي لقياس طول نفس الغرفة في اليوم التالي فيفترض أن نحصل على نفس النتيجة في ظل ثبات الظروف المحيطة بالقياس وعند ذاك سنقول أن هذه الأداة قياساتها (نتائجها) ثابتة، وإذا ما وجدنا نوعاً من الاختلاف في النتائج كل مرة فسوف لن نثق بهذه الأداة. لكن المشكلة تبدو أكثر تعقيداً عندما نستخدم أداة لقياس سمة إنسانية معينة، فلو قسنا درجة قلق فرد ما أو مجموعة أفراد في وقت معين، ثم عدنا بعد فترة زمنية أخرى فسوف يصعب علينا أن نجد أن ذلك الفرد أو تلك المجموعة تحقق نفس المستويات التي حققتها في المرة الأولى، لأن الفرد أو أي فرد من المجموعة سوف يمر بظروف معينة قد تؤدي إلى زيادة قلقه أو انخفاضه مما يؤثر سلباً على دقة ثبات ذلك المقياس، وهكذا بالنسبة للخصائص النفسية الأخرى التي يتأثر مستوى تواجدها بعوامل داخلية وطبيعية واجتماعية متنوعة، لذلك يمكن أن نكيف ما قدمنا إليه بدرجة ما فنقول أن الإختبار الثابت هو الذي يُعطي تقديرات أو قياسات متقاربة إلى حد ما عندما يُعاد تطبيقه على نفس المجموعة مرتين بينها فاصل زمني في ظل ظروف محيطية مستقرة.

ويشمل الثبات المفاهيم الآتية:

أ. الإتساق Consistancy أي اتساق أو ارتباط أو تشابه نتائج الاختبار (علاماته) مع نفسها عندما يطبق مرات متتالية.

ب. الاستقرار Stability بمعنى لو كررت عمليات قياس السمة لدى الفرد بنفس الاختبار لأظهرت علامة ذلك الفرد على الاختبار نوعاً من الاستقرار (محافظة الفرد على ترتيبه في المجموعة التي تعرضت لنفس الاختبار في مرات تطبيقه).

جـ. الموضوعية Objectivity ويقصد بها حصول الفرد على نفس العلامة أو ما يقاربها في التطبيقات اللاحقة للاختبار وإن أختلف المصححون لذلك الاختبار.

د. التكافؤ Equivalence ويعني أن مجموعتي العلامات الناشئتين عن تطبيق

الاختبار في المرتين تكونا متكافئتين (متقاربين) ويمكن الاستنتاج مما سبق وكما يؤكد (Gronlund & Linn, 199) ليست بالضرورة أن تكون قريبة منها، وإنما يكفي أن يظل الترتيب النسبي للأفراد كما هو تقريباً، لأن معاملات الارتباط التي تستخدم لإيجاد الثبات تعتمد على الترتيب النسبي.

أساليب إيجاد الثبات:

هناك أربعة أساليب شائعة الاستخدام في مجال القياس النفسي لتحديد درجة ثبات الاختبار أو المقياس. ثلاثة من تلك الأساليب تتطلب الحصول على نتائج القياس لمرتين على نفس المجموعة من الأفراد، بينما لا يتطلب الأسلوب الرابع الحصول على نتائج القياس مرتين. وفي الصفحات القادمة نستعرض بالتفصيل كيفية الحصول على معامل ثبات الاختبار بكل أسلوب من تلك الأساليب مع تطبيقات عملية واستخدام الأساليب الإحصائية الخاصة بإيجاد معامل الثبات.

(1) إعادة الإختبار Test-Retest

وفي هذا الأسلوب يتم تطبيق الاختبار المراد إيجاد معامل ثباته على مجموعة معينة من الأفراد - نفضل أن لا يقل عدد أفرادها عن ثلاثين فرداً - ثم تحسب علامات كل فرد من أفراد تلك المجموعة، وبعد مرور مدّة زمنية تتراوح من أسبوعين إلى أربعة أسابيع على الغالب يُعاد تطبيق نفس الأختبار على نفس المجموعة، وفي ظل نفس الظروف قدر الإمكان فنحصل على مجموعة ثانية من العلامات لنفس المجموعة. فإذا ما اعتبرنا مجموعة العلامات الأولى تمثل المتغير (س)، والمجموعة الثانية لعلامات نفس الأفراد تمثل المتغير (ص)، يمكننا بعدئذ تطبيق (معامل ارتباط بيرسون) الذي يمثل معامل ثبات ذلك الاختبار حيث نحكم عليه في ضوء معايير دلالة قيم معاملات الارتباط.

مثال:

أراد معلم إيجاد معامل ثبات اختبار يتكون من (20) فقرة تَمَّ تصميمه من قبله فطبقه على مجموعة من الطلاب في موعد معين ثم طبقه على نفس المجموعة بعد مرور اسبوعين على التطبيق الأولً، وكانت العلامات التي حصل عليها أفراد المجموعة في المرتين كما هي موضحه في الجدول الآتي:

الطالب	علامة الطالب في التطبيق الأول	علامة الطالب في التطبيق الثاني	إيجاد معامل الثبات
أ	15	16	تطبيق معادلة بيرسون لايجاد معامل الارتباط بين مجموعتي العلامات $ر = \frac{\text{ن مجـ س ص} - \text{مجـ س مجـ ص}}{\sqrt{[(\text{ن مجـ س}^2) - (\text{مجـ س})^2]\ [(\text{ن مجـ ص}^2 - (\text{مجـ ص})]^2}}$
ب	14	15	
ج	18	20	
د	19	14	
هـ	16	15	حيث أن :
و	12	13	ن = مجموع أفراد المجموعة
ز	11	10	س = علامات التطبيق الأول
ح	17	18	ص = علامات التطبيق الثاني
ط	14	13	ر = معامل الارتباط
ي	12	11	

ولأجل تطبيق معادلة (بيرسون) لإيجاد معامل الارتباط بين مجموعتي علامات الطلاب في التطبيقين الأول والثاني يُفترض أن نحصل على القيم التالية الضرورية لإيجاد معامل الارتباط، وهي:

مجموع قيم (س)، ومجموع قيم (ص)، ومجموع قيم $(س)^2$، ومجموع قيم $(ص)^2$، ومجموع قيم (س ص). وكما موضح في الجدول التالي:

الطلاب	س	ص	س ص	$س^2$	$ص^2$
أ	15	16	240	225	256
ب	14	15	210	196	225
ج	18	20	360	324	400
د	19	14	266	361	196
هـ	16	15	240	256	225
و	12	13	156	144	169
ز	11	10	110	121	100
ح	17	18	306	289	324
ط	14	13	182	196	169
ي	12	11	182	144	121
المجموع	148	145	2202	2256	2185

$$ر = \frac{(145x148) - (2202x10)}{\sqrt{[145x145) - (2185x10)] [(148x148) - (2256x10)]}}$$

$$ر = \frac{21460 - 22020}{\sqrt{(21025 - 2150) (21904 - 22560)}}$$

$$ر = \frac{560}{\sqrt{(825) x (656)}}$$

$$ر = \frac{560}{735.66} = 0.761$$ قيمة معامل الثبات

ومن استقراء الناتج أعلاه نستنتج أن قيمة معامل الارتباط موجبة ومقبولة، مما يعني أن هذا الاختبار يتمتع بمعامل ثبات جيد.

وعلى الرغم من أن إيجاد معامل الثبات وفقاً لهذا الأسلوب شائع الاستخدام، إلا أن هناك نقاط ضعف كثيرة تقلل من إقبال بعض الباحثين على استخدامه، ومن أهم نقاط الضعف تلك:

أ. أن بعض الأفراد أو الطلاب قد يتغير سلوكهم عند الإجابة على الاختبار في التطبيق الثاني بشكل يختلف عن سلوكهم عند الإجابة على الاختبار في التطبيق الأول وذلك بسبب الخبرة التي يكتسبها أولئك الطلاب والناشئة من تعرضهم لنفس الاختبار مرة ثانية مما يؤدي بلا شك إلى تغير واضح في علاماتهم وبالتالي تغيير في ترتيبهم في المجموع الذي حققوه في التطبيق الأول وهذا بدوره سيؤثر على معامل ثبات الاختبار ولاسيما إذا ما علمنا أن هذا التغيير لا يشمل جميع الطلاب وبمستوىً متساوٍ، بل إنه قد يشمل البعض منهم.

ب. إن بعض الطلاب يعتمدون على ذاكرتهم في تحديد الإجابات التي تم تحديدها في التطبيق الأول للاختبار وإعادتها نفسها في التطبيق الثاني مما يعني عدم اعتمادهم على التفكير والاستنتاج خلال الإجابة الثانية، لكن هذا الأمر يمكن معالجته من خلال زيادة طول الفترة الفاصلة بين التطبيقين إذا كان الاختبار يعتمد على الحفظ والتذكر كالاختبار التحصيلي المعرفي، وعدم إطالة الفترة في الاختبارات والمقاييس التي لا تعتمد على التذكر مثل مقاييس الشخصية والاتجاهات النفسية.

جـ. قد تحدث مجموعة من العوامل التي تؤثر سلباً أو إيجاباً على إجابات الطالب تُعزى إلى طول الفترة الفاصلة بين التطبيقين الأول والثاني منها عوامل النمو المعرفي أو الجسمي، ففي الجانب المعرفي يمكن أن تزداد معلومات الطالب خلال تلك الفترة عن الموضوع الذي يدور حوله الاختبار مثل المطالعات المكتبية، المعلومات التي يحصل عليها الطالب من المواد الدراسية الأخرى، مشاهدة فلم عن الموضوع أو برنامج تلفزيوني كل هذه الأمور قد تؤثر على زيادة معلومات الطالب عن موضوع الاختبار وتؤدي إلى زيادة علاماته في التطبيق الثاني.

أما بالنسبة إلى العوامل الجسمية فتتمثل في التغيرات البيولوجية التي قد تطرأ على الطالب خلال الفترة الفاصلة بين التطبيقين، مثال ذلك إنتقال الطالب من مرحلة الطفولة المتأخرة إلى مرحلة المراهقة، أو من المراهقة إلى مرحلة الرشد وبالتالي إحتمال تأثير هذا النمو على درجات الأفراد في التطبيق الثاني بما يؤدي إلى تأثر معامل الثبات بذلك. وأخيراً لابد من الإشارة إلى أن هناك صعوبة أخرى تواجه الباحث في اتباع هذا الأسلوب تتمثل في صعوبة جمع أفراد عينة التطبيق مرتين لاسيما في المؤسسات الإنتاجية حيث يؤثر ذلك سلباً على العملية الإنتاجية وصعوبة موافقة المسؤولين على هذا الإجراء.

بأختصار، تتلخص مشاكل طريقة اعادة تطبيق الاختبار في المدة الزمنية التي تفصل بين التطبيقين، فعندما تكون قصيرة تبرز مشكلة الذاكرة، وحينما تطول تبرز مشكلة النمو. وحلاً لهذه المشاكل اقترح العاملون في مجال القياس طريقة الصور المتكافئة.

(2) الصور المتكافئة Equivalent Forms

يختصر هذا الأسلوب العديد من المشاكل التي تمَّ التطرق إليها عند استخدام إسلوب (إعادة الاختبار) لإيجاد معامل الثبات. ففي طريقة الصور المتكافئة، يعمد الباحث إلى صياغة صورتين لنفس الاختبار، تتشابهان في الهدف الذي نريد تحقيقه أو السلوك الذي نريد قياسه بالاختبار وكذلك نحرص على أن يكون محتوى الاختبار في الصورتين مغطىً بنفس درجة الشمول، يضاف إلى ذلك أن من الضروري أن يكون مستوى صعوبة الفقرات لاسيما في الاختبارات التحصيلية متساوياً في الصورتين. ويقترح كل من (الكناني وجابر، 1995) شروط يجب توافرها في هاتين الصورتين منها:

تساوي متوسطي الأداء على الصورتين أثناء تجريبهما على عينة التقنين، وتساوي الانحرافات المعيارية، وتساوي عدد الفقرات، وتساوي طول الاختبار وتشابه الصورتين في إجراءات التطبيق والتصحيح والتوقيت.

وإذا ما أكمل الباحث صورتي الاختبار وفقاً للمعايير السابقة الذكر، يمكن

تطبيقهما معاً أي إعطائهما لكل فرد ليجيب عنهما مرةً واحدة، وهناك من يقترح إعطاء الصورة الأولى للطلاب، ثم أخذ استراحة بسيطة وإعطاء الصورة الثانية بعد الاستراحة مباشرةً. أما كيفية إيجاد معامل ثبات الاختبار وفقاً لهذا الأسلوب فتتم وفقاً للخطوات التالية:

1- يتم تصحيح كل صورة من الاختبار لكل طالب على حده، فنحصل على علامتين لكل طالب إحداهما تمثل علامته على الصورة (أ) والأخرى تمثل علامته على الصورة (ب).

2- تمثل علامات الصورة (أ) قيم (س)، وتمثل علامات الصورة (ب) قيم (ص) كمتغيرين نريد إيجاد العلاقة الارتباطية بينهما.

3- نختار معامل الارتباط المناسب وفقاً لنوع المتغيرات، فإذا كانت علامات الطالب من النوع المتصل يمكن تطبيق (معامل ارتباط بيرسون)، وكما في المثال التالي:

مثال:

الطالب	علامة الطالب على الصورة الأولى	علامة الطالب على الصورة الثانية
أ	15	16
ب	17	19
ج	13	15
د	14	15
هـ	18	19
و	11	12
ز	14	16

ومن خلال تطبيق معامل ارتباط بيرسون نجد أن قيمة معامل الارتباط بين مجموعتي العلامات يساوي (0.976) وتشير هذه القيمة إلى معامل ثبات هذا الاختبار وتسمى أيضاً (معامل التكافؤ).

ومما تجدر الإشارة له أن هذا الأسلوب وعلى الرغم من أنه تفادى نقاط الضعف التي تم تحديدها في أسلوب (إعادة الاختبار) لاسيما الناشئة عن الفترة الزمنية الفاصلة بين التطبيقين، وكذلك المشاكل الناشئة عن جمع أفراد العينة مرتين إلا إنه يحمل عدداً من الصعوبات التي تقف في طريق تطبيقه منها: الجهد الكبير المبذول في تصميم صورتين لنفس الاختبار يتشابهان في خصائص عديدة، وكذلك صعوبة تحقيق تكافؤ كامل بين الصورتين مهما بذل الباحث من جهودٍ لتحقيق هذا الهدف.

(3) التجزئة النصفية Split half

في هذا الأسلوب يمكن تلافي نقاط الضعف الموجودة في الأسلوبين السابقين، إعادة الاختبار والصور المتكافئة حيث لا نحتاج إلى تجميع المفحوصين مرتين وما ينشأ عن ذلك من مشاكل بسبب الفاصل الزمني بين التطبيقين، أو الصعوبات التي يتحملها الباحث من جراء إيجاد صورتين لنفس الاختبار، ففي إسلوب التجزئة النصفية يقوم الباحث بتطبيق الاختبار مرةً واحدة على مجموعة المفحوصين وبعد إكمالهم للإجابات يتم تصحيحها، ثم توزع علامات الأسئلة إلى مجموعتين هما: مجموعة علامات الأسئلة الفردية لتمثل المتغير (س)، ومجموعة علامات الأسئلة الزوجية لتمثل المتغير (ص)، ثم نوجد العلاقة الارتباطية بين علامات النصفين.

وهناك من يقسم الاختبار إلى قسمين بطريقة أخرى حيث يتضمن النصف الأول مجموع علامات الأسئلة من الرقم (1) إلى منتصف المجموع الكلي، ثم يجمع علامات أسئلة النصف الثاني، وبذلك يكون لكل فرد في المجموعة علامتين هما (س) و (ص) ثم توجد العلاقة بين مجموعتي العلامات.

وهناك أسلوب ثالث يعتمد على سحب أرقام الأسئلة بشكل عشوائي ويوزعها على مجموعتين تمثلان نصفي الاختبار ثم يُوجِد العلاقة بينهما. ويمكن القول أن أسلوب تجزئة الاختبار إلى نصفين وفقاً للأرقام الفردية والزوجية لأسئلة الاختبار هي الأفضل حيث يمكن أن يوزع المفحوص جهده

ووقته بنفس القدر بين نصفي الاختبار وخاصةً إذا كانت وحداته متدرِّجة من السهولة إلى الصعوبة، وفيها أيضاً ضمان لثبات العوامل المؤثرة على نتائج القياس في النصفين بينما لا تتوافر هذه الميزات في طريقة قسمة الاختبار إلى نصفين أول وثاني على التوالي، أو توزيع الأسئلة عشوائياً وفقاً لأرقامها على النصفين الأول والثاني حيث يمكن أن تؤثر عوامل التعب والملل التي يعاني منها المفحوص على إجابته على أسئلة النصف الثاني، إضافة إلى احتمال أن تكون أسئلة أحد النصفين أصعب أو أسهل من أسئلة النصف الثاني. وبعد أن يحدد الباحث علامات كل طالب على النصفين الأول والثاني ويوجد معامل الارتباط بينهما. ويجب أن لا يغيب عن باله أن هذا المعامل سيكون منخفضاً لأنه أدى إلى اختزال طول الاختبار إلى النصف، لذلك أوجب الباحثون ضرورة استخدام معادلة سبيرمان - براون لتصحيح معامل ثبات التجزئة النصفية.

معادلة سبيرمان - براون Sperman-Brown

ويشيع استخدام هذه المعادلة بصورة خاصة في اختبارات التحصيل والقدرات، ونصها:

$$ر_{1.1} = \frac{2 \times ر_{\frac{1}{2}.\frac{1}{2}}}{1 + ر_{\frac{1}{2}.\frac{1}{2}}}$$

حيث أن:

$ر_{1.1}$ = معامل ثبات الاختبار الكلي.

$ر_{\frac{1}{2}.\frac{1}{2}}$ = معامل ثبات نصفي الاختبار.

ويمكن توضيح كيفية إيجاد معامل ثبات الاختبار بطريقة التجزئة النصفية (زوجي- فردي) في المثال التالي:

مثال:

أراد باحث أن يوجد معامل ثبات اختبار يتكون من (12) سؤال يجيب عليها الطالب بإحدى كلمتي (صح) أو (خطأ) ليحصل على إحدى علامتي (1) أو (صفر). طبق الاختبار على عشرة طلاب وكانت علاماتهم على تلك الأسئلة كما يأتي:

الطلبة	فقرات الاختبار												مجموع الفردية (س)	مجموع الزوجية (ص)
	1	2	3	4	5	6	7	8	9	10	11	12		
أ	1	1	1	0	0	1	1	1	1	0	0	1	4	4
ب	0	1	1	1	1	1	1	0	1	1	0	1	4	5
ج	1	1	1	1	1	1	1	1	1	1	1	1	6	6
د	0	0	1	1	1	0	0	1	1	1	1	1	4	4
هـ	1	1	1	0	0	1	1	1	1	1	0	1	4	5
و	1	1	1	1	0	0	1	1	1	0	1	1	5	4
ز	0	0	0	1	1	0	0	1	1	1	0	0	2	3
ح	1	1	1	1	1	1	1	1	1	1	1	1	6	6
ط	1	0	0	1	1	1	0	0	1	1	0	0	3	3
ي	1	1	1	1	1	1	1	1	1	1	1	1	6	6

الحل:

1- يتم تجميع ما حصل عليه كل طالب من علامات على الأسئلة الفردية ذات التسلسل (1، 3، 5، 7، 9، 11) ووضعها تحت حقل المتغير (س).

2- يتم تجميع ما حصل عليه كل طالب على الأسئلة الزوجية ذات التسلسل (2، 4، 6، 8، 10، 12) ووضعها تحت حقل المتغير (ص).

3- نطبق معادلة (بيرسون) لإيجاد العلاقة بين علامات الأسئلة الفردية (س)، وعلامات الأسئلة الزوجية (ص) لنحصل على قيمة العلاقة بين المتغيرين.

4- نظراً لأن قيمة معامل الارتباط الناتجة عن الخطوات السابقة تنتج عنها معامل ارتباط بين نصفي الاختبار وليس اختبارين سيكون مخفضاً بلاشك، لذلك علينا أن نطبق (معادلة سبيرمان - براون) لنحصل على معامل ثبات الاختبار.

يتطلب إيجاد معامل الثبات وفقاً لمعامل ارتباط بيرسون تحديد قيم (مجموع س، مجموع ص، مجموع س ص، مجموع $س^2$، مجموع $ص^2$) وكما مدرج في الجدول التالي:

الطلبة	س	ص	س ص	$س^2$	$ص^2$
أ	4	4	16	16	16
ب	4	5	20	16	25
ج	6	6	36	36	36
د	4	4	16	16	16
هـ	4	5	20	16	225
و	5	4	20	25	16
ز	2	3	6	04	09
ح	6	6	36	36	36
ط	3	3	9	09	09
ي	6	6	36	36	36
المجموع	44	46	215	210	224
	مجـ س	مجـ ص	مجـ س ص	مجـ $س^2$	مجـ $ص^2$

$$ر = \frac{46x44 - 215x10}{\sqrt{[46x46 - 224x10]\,[44x44 - 210x10]}}$$

$$ر = \frac{2024 - 2150}{\sqrt{[2116 - 2240]\ [1936 - 2100]}}$$

$$ر = \frac{126}{\sqrt{(124) \times (164)}}$$

$$ر = \frac{126}{142.60} = 0.88$$ قيمة معامل الارتباط بين نصفي الاختبار

وبعد تطبيق معادلة سبيرمان - براون نحصل على معامل ثبات الاختبار

$$ر_{1.1} = \frac{0.88 \times 2}{0.88 + 1} = \frac{1.76}{1.88}$$

$ر_{1.1} = 0.936$ وهو معامل ثبات مرتفع وموجب.

معادلة رولون Rulon

وفي هذه المعادلة التي تستخدم في التجزئة النصفية يتم الاعتماد على الأساس النظري الذي يفترض أن التباين الكلي لدرجات الاختبار يمكن أن يقسم إلى قسمين: قسم ناتج عن الخطأ مضاف إليه القسم الآخر الناتج عن الفروق الفردية الحقيقية، ونص المعادلة هو:

$$ر = 1 - \frac{ع^2ف}{ع^2_ك}$$

حيث أن:

$ع^2ف$ = تباين الفروق بين درجات نصفي الاختبار ويمثل الخطأ التجريبي أي من خلال طرح (س – ص).

$ع^2$ = تباين درجات الاختبار الكلي. بعد ضم (س + ص).

مثال:

ومن بيانات المثال السابق يمكننا الحصول على معامل ثبات الاختبار وفقاً لهذه المعادلة باتباع الخطوات التالية:

1- نطرح كل قيمة (ص) من قيمة (س) لنحصل على قيم (س - ص). ونجد تباينها ليمثل تباين الفروق بين نصفي الاختبار .

2- نجمع كل قيمة في المتغير (س) مع ما يقابلها من قيم المتغير (ص) لنحصل على قيم (س+ص). ونجد تباين هذه القيم ($ع^2_ك$) الذي يمثل التباين الكلي للاختبار ($ع^2_ك$).

ولإيجاد تباين القيم هناك طريقتان الأولى أن نجد متوسط تلك القيم، ثم نطرح هذا المتوسط من كل قيمة من تلك القيم (س - $\overline{س}$)

الطلبة	س	ص	س - ص	س + ص
أ	4	4	0	8
ب	4	5	1-	9
ج	6	6	0	12
د	4	4	0	8
هـ	4	5	1-	9
و	5	4	1	9
ز	2	3	1-	5
ح	6	6	0	12
ط	3	3	0	6
ي	6	6	0	12

ثم نربع الخطوة السابقة، أي إيجاد قيمة $(س-\overline{س})^2$، وبعدها نجد حاصل جمع القيم المربعة أي مجموع $(س-\overline{س})^2$ ونقسم الناتج على عدد القيم (ن).

ومن مثال السابق يمكن أن نجد التباين الكلي للاختبار كما يأتي:

القيم (س)	س - $\bar{س}$	$(س - \bar{س})^2$
8	1-	1
9	0	0
12	3	9
8	1-	1
9	0	0
9	0	0
5	4-	16
12	3	9
6	3-	9
12	3	9
مجـ س = 90		مجـ = 54

$$\text{الوسط الحسابي للقيم} = \frac{\text{مجموع القيم}}{\text{عدد القيم}}$$

$$\bar{س} = \frac{\text{مجـ س}}{ن} = \frac{90}{10} = 9 \text{ الوسط الحسابي.}$$

$$ع^2 = \frac{\text{مجـ}(س - \bar{س})^2}{ن} = \frac{54}{10} = 5.4 \text{ تباين الاختبار}$$

أما الطريقة الثانية لإيجاد التباين فهي أبسط من الطريقة الأولى وفيها يتم تجاوز مشكلة الكسور العشرية الناشئة عن طرح الوسط الحسابي من كل قيمة وتتلخص في إيجاد حاصل جمع القيم أولاً (مجـ س) ثم تربيع كل قيمة من القيم ثم إيجاد مجموع مربعات القيم ثم تطبيق معادلة التباين كالآتي:

س	س^2
8	64
9	81
12	144
8	64
9	81
9	81
5	25
12	144
6	36
12	144
90	864

$$ع^2 = \frac{\text{ن مجـ س}^2 - (\text{مجـ س})^2}{\text{ن}^2}$$

$$ع^2 = \frac{90x90 - 864x10}{10x10}$$

$$ع^2 = \frac{54010}{0}$$

ع^2 = 5.4 تباين الاختبار

ونلاحظ أننا توصلنا إلى نفس قيمة التباين في الطريقتين الأولى والثانية، ولأجل التوصل إلى إيجاد معامل الثبات بطريقة رولون علينا أن نجد تباين الفروق بين درجات نصفي الاختبار وهي القيم المدرجة في العمود الثالث من الجدول الأساسي أي قيم (س-ص) وكما يأتي:

س - ص	س^2
0	0
1-	1
0	0
0	0

$$ع^2 = \frac{\text{ن مجـ س}^2 - (\text{مجـ س})^2}{\text{ن}^2}$$

$$ع^2 = \frac{(2x2) - (4x10)}{10x10}$$

س - ص	$س^2$
1-	1
1	1
1-	1
0	0
0	0
0	0
(مجـ س) = 2	(مجـ $س^2$) = 4

$$ع^2 = \frac{36}{100}$$

$ع^2 = 0.36$ تباين الاختبار

3 - نطبق معادلة رولون وهي:

$$ر = 1 - \frac{ع^2_{ف}}{ع^2_{ك}}$$

$$ر = 1 - \frac{0.36}{5.4}$$

ر = 1 - 0.067 = 0.93 معامل الثبات

ويلاحظ أن قيمة معامل الثبات موجبة وعالية ويمكن الوثوق بثبات هذا الاختبار، كما يلاحظ أن هذه القيمة هي مقاربة جداً لقيمة معامل الثبات الذي تم إيجاده بالطريقة الأولى أي بعد استخدام معادلة (سبيرمان - براون) ص264.

معادلة جتمان Guttman

تستخدم هذه المعادلة عندما يلاحظ الباحث عدم تساوي تباين علامات الطلاب على نصفي الاختبار وتتميز ببساطة الأسلوب المستخدم لإيجاد معامل ثبات الاختبار وتتطلب إيجاد التباين لكلا النصفين وكذلك إيجاد التباين الكلي للاختبار، ونص المعادلة هو:

$$\text{معامل ثبات الاختبار} = 2 \left(1 - \frac{ع_1^2 + ع_2^2}{ع_ك^2}\right)$$

حيث أن: $ع_1^2$ تباين النصف الأول للاختبار.

$ع_2^2$ تباين النصف الثاني للاختبار.

$ع_ك^2$ التباين الكلي للاختبار

مثال:

من بيانات المثال السابق يمكن إيجاد معامل الثبات حسب معادلة جتمان كما يأتي:

1- نجد تباين علامات الطلاب على النصف الأول من الاختبار:

علامات الطلاب على النصف الأول س	مربع علامات النصف الأول $س^2$
4	16
4	16
6	36
4	16
4	16
5	25
2	4
6	36
3	9
6	36
(مجـ س) = 44	(مجـ $س^2$) = 210

$$ع_1^2 = \frac{ن \text{ مجـ } س^2 - (\text{مجـ } س)^2}{ن^2}$$

$$ع_1^2 = \frac{(44x44) - (210x10)}{10x10}$$

$$ع_1^2 = \frac{164}{100}$$

$ع_1^2 = 1.64$ تباين النصف الأول

2- نجد تباين علامات الطلاب على النصف الثاني من الاختبار:

علامات الطلاب على النصف الثاني س	مربع علامات النصف الثاني $س^2$
4	16
5	25
6	36
4	16
5	25
4	16
3	9
6	36
3	9
6	36
(مجـ س) = 46	(مجـ $س^2$) = 224

$$ع_2^2 = \frac{ن \text{ مجـ } س^2 - (\text{مجـ } س)^2}{ن^2}$$

$$ع_1^2 = \frac{(46x46) - (224x10)}{10x10}$$

$$ع_2^2 = \frac{124}{100}$$

$ع_2^2 = 1.24$ تباين النصف الثاني

3- بعد حصولنا على تباين نصفي الاختبار نجد التباين الكلي للاختبار وهو موجود أصلاً في الصفحات السابقة وقيمته (5.4)، بعدها نطبق معادلة جتمان لنحصل على معامل ثبات الاختبار:

$$ر = 2 \left(1 - \frac{ع_1^2 + ع_2^2}{ع_ك^2}\right)$$

$$ر = 2 \left(1 - \frac{1.24 + 1.64}{5.4}\right) \qquad ر = 2 \left(1 - \frac{2.88}{5.4}\right)$$

ر = 2 (1 - 0.53) = 2 × 0.47 = 0.93 معامل ثبات الاختبار

ويلاحَظ أن معامل ثبات الاختبار موجب وعالي، كما أنه لا يختلف عن قيمة معامل الثبات الذي تمَّ إيجاده بطريقة رولون.

جـ. معادلة هورَّست Horest

تستخدم هذه المعادلة عندما يكون جزئي الاختبار غير متساويين من حيث عدد فقراتهما، كأن يكون عدد الفقرات الفردية أكبر من عدد الفقرات الزوجية، أو أن تكون فقرات الجزء الأول أكبر أو أقل عدداً من فقرات الجزء الثاني.

ونص المعادلة هو:

$$\text{ر}_{\text{ك}} = \frac{\text{ر}_{2.1}\sqrt{[\text{ر}^{2}_{2.1} + 4\,\text{أ ب}\,(1 - \text{ر}^{2}_{2.1}) - \text{ر}_{2.1}]}}{2\,\text{أ} \times \text{ب}\,(\text{ر}_{2.1}\ 2\,\text{أ ب})}$$

حيث أن:

$\text{ر}_{\text{ك}}$ = معامل ثبات الاختبار ذو الجزئين غير المتساويين.

أ = نسبة عدد فقرات الجزء الأول (الفردي) إلى عدد فقرات الاختبار كله.

ب = نسبة عدد فقرات الجزء الثاني (الزوجي) إلى عدد فقرات الاختبار كله.

$\text{ر}_{2.1}$= معامل ارتباط جزئي الاختبار.

مثال:

لو أن الاختبار الذي تم تصميمه يتكون من (25 فقرة) منها (13) فقرة ذات تسلسل فردي، و(12) فقرة ذات تسلسل زوجي وكانت علامات مجموعة من الطلبة على كل من الفقرات الفردية والفقرات الزوجية كما يأتي:

الطالب	مجموع علامات الطلاب على الفقرات الفردية (س)	مجموع علامات الطلاب على الفقرات الزوجية (ص)	العلامة الكلية للطالب
أ	12	11	23
ب	10	10	20
ج	13	11	24
د	9	8	17

الطالب	مجموع علامات الطلاب على الفقرات الفردية (س)	مجموع علامات الطلاب على الفقرات الزوجية (ص)	العلامة الكلية للطالب
هـ	7	6	13
و	10	9	18
ز	13	12	25
ح	11	10	21
ط	8	9	17
ي	12	11	23

الحل: لإيجاد معامل الثبات بين النصفين غير المتساويين في عدد فقراتهما وفقاً لمعادلة (هورست) نتبع الخطوات التالية:

1- نوجد معامل الارتباط بين علامات نصفي الاختبار من خلال تطبيق معامل ارتباط بيرسون حيث بلغت قيمته (0.93)

2- نجد نسبة عدد الفقرات الفردية إلى المجموع الكلي لفقرات الاختبار وتساوي: $\frac{13}{25} = 0.52$، وتمثل قيمة (أ) في معادلة هورست.

3- نجد نسبة عدد الفقرات الزوجية إلى المجموع الكلي لفقرات الاختبار وتساوي: $\frac{12}{25} = 0.48$، وتمثل قيمة (ب) في معادلة هورست.

4- نطبق معادلة هورست بتعويض رموزها لنحصل على معامل ثبات الاختبار.

$$ر_ك = \frac{[0.96 - ({}^2(0.96) - 1)\,(0.48\text{x}0.52)4 \;+\; {}^2(0.96)]\sqrt{}\; 0.96}{(0.48\text{x}0.52)2 \text{ x } (0.96) \text{ x } (0.48) \text{ x } (0.52\text{x}2)}$$

$$ر_ك = \frac{0.04\sqrt{}\;\; 0.96}{(0.478 \text{ x } 0.48) \text{ x } 1.04}$$

رك = 0.81 معامل ثبات الاختبار ذو الجزئين غير المتساويين، وهو موجب ومرتفع.

(4) إيجاد معامل الثبات باستخدام تحليل التباين Analysis of Variance

إن إيجاد معامل الثبات من خلال استخدام أسلوب (تحليل التباين) يقوم على فكرة مقاربة لفكرة الاتساق الداخلي Internal Consistercy حيث يهدف إلى التأكد من إتساق أداء الفراد على عموم الاختبار من فقرة إلى أخرى، كما يدل التجانس الكلي لفقرات الاختبار على استقرار استجابات الأفراد على الاختبار، ويشير ويرسمى وجرس (Wiersma and Jurs, 1990) إلى أن استخدام المصطلح (الداخلي Internal) بسبب أن جميع الفقرات تشترك في إعطاء صفة الأصالة للاختبار.

وعندما تكون العوامل الأخرى متساوية، فإن محتوى الاختبار كلما كان متشابهاً فإن ثبات الإتساق الداخلي سيكون مرتفعاً.

ومن أهم المعادلات المستخدمة في هذا الأسلوب هي:

أ- معادلة كودر - ريتشاردسون (20) Kuder - Richardson Formula - 20

وتستخدم هذه المعادلة على وجه الخصوص مع الاختبارات ذات الاختيارات التي يكون تصحيح فقراتها بصفر أو واحد كالاختبارات التي تكون إجاباتها الصحيحة واحدة فقط من اختيارين، أو ثلاثة أو أربعة أو خمسة وفي هذه المعادلة لا يجزأ الاختبار إلى جزئين كما في حالة التجزئة النصفية بل إلى (K) أي بعدد الفقرات التي يتكون منها الاختبار ونص هذه المعادلة هو:

$$r = \frac{n}{n - 1}\left[1 - \frac{\Sigma pq}{S^2}\right]$$

وتكتب باللغة العربية على الشكل الآتي:

$$ر = \frac{ن}{ن - 1}\left[1 - \frac{مجـ ص خ}{ع_ك^2}\right]$$

حيث أن:

ر= معامل ثبات الاختبار.

ن= عدد فقرات الاختبار.

ص = نسبة عدد الطلاب الذين أجابوا إجابة صحيحة على الفقرة إلى العدد الكلي للطلبة في المجموعة.

خ = نسبة عدد الطلاب الذين فشلوا في تقديم الإجابة الصحيحة على الفقرة إلى العدد الكلي للطلبة في المجموعة

$ع^2_ك$ = قيمة التباين لعموم الاختبار.

مثال:

فيما يلي علامات عشرة من الطلاب على اختبار يتكون من ثمان فقرات:

الطلبة	الفقرات 1	2	3	4	5	6	7	8	مجموع العلامات (س)	$س^2$
أ	0	1	1	1	1	0	0	1	6	36
ب	1	1	1	0	1	1	0	1	4	16
ج	1	1	1	1	1	1	1	1	6	36
د	1	0	0	1	1	1	1	1	3	09
هـ	0	1	1	1	1	1	0	1	6	36
و	0	0	1	1	1	0	1	1	8	64
ز	1	0	0	1	1	1	0	0	7	49
ح	1	1	1	1	1	1	1	1	4	16
ط	1	1	0	0	1	1	0	0	5	25
ي	1	0	1	0	1	0	1	0	4	16
نسبة الصواب (ص)	0.9	0.8	0.5	0.4	0.6	0.5	0.7	0.1	مجـ س = 53	مجـ س = 303
نسبة الخطأ (خ)	0.1	0.2	0.5	0.6	0.4	0.5	0.3	0.1		

مجـ ص خ $= + (0.1 \times 0.9) + (0.2 \times 0.8) + (0.5 \times 0.5) + (0.6 \times 0.4)$

$(0.4 \times 0.6) + (0.5 \times 0.5) + (0.3 \times 0.7) + (0.1 \times 0.9) = 1.53$

من الجدول السابق:

1- يتم حساب نسبة الطلاب الناجحين لكل فقرة (ص) ونسبة الطلاب الراسبين بكل فقرة (خ) وذلك بقسمة (مجموع ص) على العدد الكلي للممتحنين، وكذلك قسمة (مجموع س) على العدد الكلي للممتحنين وهي عادة تستخرج من خلال (1-ص).

ويمكن ملاحظة النسب الخاصة بكل من (ص) و(خ) في الجزء الأسفل من الجدول السابق.

2- نوجد تباين الاختبار $ع_ك^2$ من خلال مجموع العلامات التي حصل عليها كل طالب لتمثل قيم (خ) ثم نجمع العلامات الكلية لعموم الطلبة لتمثل (مجـ س) ثم نربع كل قيمة من قيم (خ) لنحصل على قيمة (مجـ $س^2$) ونوجد التباين بتطبيق المعادلة التالية:

$$ع_ك^2 = \frac{ن\ مجـ\ س^2 - (س)^2}{ن^2}$$

$$ع_ك^2 = \frac{10 \times 303 - 53 \times 53}{100}$$

$$ع_ك^2 = \frac{3030 - 2809}{100}$$

$$ع_ك^2 = \frac{221}{100} \qquad ع_ك^2 = 2.21 \text{ التباين الكلي للاختبار}$$

3- نوجد قيمة مجـ ص(1-ص) من خلال جمع نواتج حاصل ضرب كل قيمة من (ص) في (س) أي (مجـ ص خ)، وكما موجود في أسفل الجدول وتساوي (1.53).

4- نعوض رموز المعادلة بالأرقام فنحصل على:

$$ر = \frac{8}{8-1}\left[1 - \frac{1.53}{2.21}\right]$$

$$ر = 1.14 \times 0.31 \qquad ر = 0.35 \text{ معامل الثبات}$$

ب- معادلة كودر ريتشاردسون (21) **Kuder - Richardson Formula - 21**

ويرمز لها بالرقم (21) لتمييزها عن المعادلة السابقة (20) وهي مفيدة للاستخدام في ايجاد ثبات الاختبارات الموضوعية، وتحتاج الى حسابات أبسط من المعادلة الأولى فهي تتطلب ايجاد الوسط الحسابي والتباين لمجموعة العلامات الخاصة بالطلاب المتحنين وعدد فقرات الاختيار ثم نعوض هذه القيم الثلاثة برموز المعادلة وهي:

$$ر_{س ص} = \frac{ن}{ن - 1}\left[1 - \frac{\bar{س}\,(ن - \bar{س})}{ع_{ك}^{2} \times ن}\right]$$

حيث أن:

ن = عدد فقرات الاختبار.

سَ = الوسط الحسابي لعلامات الطلاب.

ع = التباين الكلي لعلامات الطلاب.

ويذكر كل من مهرنزوليهمان (Mehrrens and Lehmann, 1991) الفروق بين المعادلة (KR-20) وهذه المعادلة (KR21) في أن الأخيرة تفترض أن جميع فقرات الاختبار متساوية في صعوبتها، فإذا لم يتحقق هذا الإفتراض فإن (KR-21) سينتج عن استخدامها تقديراً منخفضاً لمعامل الثبات، ولذلك لا يشيع استخدام المعادلتين من قبل مصممي اختبارات التحصيل.

مثال: من خلال البيانات الواردة في المثال السابق يمكن ايجاد الوسط الحسابي ومقداره (5.3) وتباين درجات الطلاب ومقداره (2.21) وبذلك يكون معامل ثبات ذلك الاختيار وفقاً لهذه المعادلة كالآتي:

$$= \frac{8}{8 - 1}\left[1 - \frac{5.3\,(8 - 5.3)}{2.21 \times 8}\right]$$

$$= 1.14 \left[1 - \frac{14.31}{17.68}\right] = 0.22$$

- وهو قريب من معامل الثبات الذي تم ايجاده وفقاً للمعادلة. KR-20

ج- معادلة الفا - كرونباخ **Coefficient Alpha**

لقد تم تطوير هذه المعادلة من قبل كرونباخ عام (1951) وهي تعميم لمعادلة (KR-20) عندما يكون تصحيح فقرات الاختبار ليس ثنائياً أي لا يأخذ (صفر أو واحد) بل يمكن أن تأخذ الفقرات درجات (1، 2، 3. . الخ) وهذه المعادلة هي في الأساس تشبه معادلة (KR-20) ما عدا الجزء الخاص بالمعادلة المذكورة وهو (Σpq) أي مجموع حاصل ضرب نسبة الصعوبة في نسبة السهولة لكل فقره من فقرات الاختبار، حيث تم استبدال ذلك الجزء بشكل آخر هو (ΣS^2) أي (مجـ ع^2) الذي يعني مجموع تباينات علامات الطلاب على كل فقرة من الفقرات. ويذكر كل من مهرنس وليهمان Meherebs & Lehman, 1991) أن هذه المعادلة قابلة للاستخدام في الاختبارات المقالية التي تأخذ اجاباتها مدىً من القيم يمتد الى أكثر من قيمتين، وكذلك اختبارات الشخصية والاتجاهات التي تمتد فيها الاستجابات الى ثلاث أو خمس أو أكثر مثل (موافق جداً - موافق - لا أدري - غير موافق - غير موافق جداً).

إن الفكرة الأساسية لهذه المعادلة تقوم على حساب الإرتباطات بين العلامات لمجموعات الثبات على جميع الفقرات التي يتكون منها الاختبار، وكأننا نقسم الاختبار ليس إلى جزئين كما في طريقة التجزئة النصفية، بل إلى عدد من الأجزاء هو عدد فقرات الاختبار وكأن كل فقرة تمثل إختباراً فرعياً له درجات تمثل درجات الطلبة على تلك الفقرة، فإذا كان عدد فقرات الاختبار على سبيل المثال (25 فقره) فان مجموع الارتباطات الممكن إيجادها هو:

$$\frac{\text{ن}\ (\text{ن} - 1)}{2} \quad \text{ويساوي} \quad \frac{25\ (25 - 1)}{2} = 300 \ \text{معامل ارتباط}.$$

ويرمز لمعادلة ألفا - كرونباخ بالآتي:

$$\alpha = \frac{\text{ن}}{\text{ن} - 1}\left[1 - \frac{\text{مجـ ع}^2\text{ف}}{\text{ع}^2\text{ك}}\right]$$

حيث أن: ن = عدد فقرات الاختبار

ع^{2}ف = تباين درجات الطلاب على كل فقره من الفقرات.

ع^2 ك = تباين العلامات الكلية للطلبة على الاختبار.

مثال:

فيما يلي استجابات خمسة طلاب على مقياس إفتراضي للاتجاهات نحو المواد الدراسية التربوية في كلية التربية مكون من ست فقرات يسستجيب لها الطالب بإحدى الاستجابات الثلاثة: موافق، لا أدري، غير موافق. حيث تعطى علامات (3، 2، 1 في حالة الفقرة الموجبة)، و(1، 2، 3 في حالة الفقره السالبه). المطلوب ايجاد معامل الفا لهذا المقياس.

الطلاب	الفقرات					
	1	2	3	4	5	6
أ	2	2	3	3	2	2
ب	2	3	3	2	1	1
ج	3	3	3	1	2	3
د	2	1	3	1	2	1
هـ	3	3	1	1	1	2
$ع^2ف$	0.24	0.64	0.64	0.64	0.24	0.14

لإيجاد معامل ألفا يفترض ايجاد قيم تباين كل فقره على حده ويمكن التوصل الى ذلك من خلال تحديد اجابات جميع أفراد المجموعة على كل فقرة، واعتبارها (س) ثم تجمع تلك القيم لتحصل على (مجـ س) ثم نربع كل قيمة من قيم (س) لنحصل على ($س^2$) ثم مجموع (مجـ $س^2$) ونطبق معادلة التباين وهي:

$$ع^2 = \frac{ن \text{ مجـ } س^2 - (\text{مجـ } س)^2}{ن^2}$$

ولنبدأ بالفقرة (1):

س	$س^2$
2	4
2	4
3	9
2	4
3	9
12	30
مجـ س	مجـ $س^2$

$$ع^2_ك = \frac{(12x12) - 30 \times 5}{5 \times 5} = \frac{144 - 150}{25}$$

= 0.24 تباين الفقرة الأولى

وبنفس الطريقة نجد تباين الفقرة الثانية = 0.64 وتباين الفقرة الثالثة = 0.64، وتباين الفقرة الرابعة = 0.64 وتباين الفقرة الخامسة= 0.24 وتباين الفقرة السادسة = 0.14 وباضافة تلك التباينات لبعضها نحصل على قيمة ($\sum$ع^{2}ف) = 0.24 + 0.64 + 0.64 + 0.64 + 0.24 + 0.14 = 2.54 أما الخطوة الثانية فتتطلب ايجاد قيمة تباين الدرجات الكلية للطلاب على المقياس، أي مجموع ما حصل عليه كل طالب على جميع فقرات المقياس واعتبارها قيم (س) ثم نربعها لنحصل على (س^2)، ثم نجمع قيم (س) وقيم (س^2)، ونطبق معادلة ايجاد التباين الموضحة أعلاه.

الطلاب	س	$س^2$
أ	14	196
ب	12	144
ج	15	225
د	10	100
هـ	11	121
	62	786
	مجـ س	مجـ $س^2$

$$ع^2ك = \frac{(62x62) - 786x5}{5 \times 5} = \frac{3844 - 3930}{25}$$

= 3.44 التباين الكلي للمقياس

وبعد تطبيق معادلة الفا كرونباخ نحصل على معامل ألفا كالآتي:

$$\alpha = \frac{ن}{ن - 1}\left[1 - \frac{مجـ ع^2 ف}{ع^2 ك}\right]$$

$$= \frac{6}{6 - 1}\left[1 - \frac{2.54}{3.44}\right]$$

$= 1.2\ (1 - 0.74) = 1.2 \times 0.26 = 0.31$ قيمة معامل ألفا كرونباخ.

وبعد استعراض الأساليب الشائعة الاستخدام لايجاد ثبات أدوات القياس لا بد من الاشارة الى أنه لا يوجد أسلوب نموذجي يمكن تعميمه أو تفضيله على الأساليب الأخرى لأن لكل حالة نريد قياسها بأداة معينة تتأثر بظروف وعوامل متنوعة سنأتي عليها لاحقاً هي التي تحدد الأسلوب المفروض استخدامه وما اذا كان يتناسب مع الغرض الذي تؤديه تلك الأداة. فإختبارات السرعة مثلاً يُناسبها (معامل التكافؤ) ولا يناسبها (معامل التجانس)، والاختبارات التي يصممها معلم الصف قد لا يُفيد معها (معامل إعادة الاختبار) للأسباب التي سبق ذكرها عند الحديث عن طريقة إعادة التطبيق بل إن الأسلوب المناسب لها هو تطبيق الاختبار لمرة واحدة ، لذلك على الباحث أن يعرف أنه ليس المهم حساب معامل الثبات، بل المهم إختيار طريقة إيجاد الثبات التي تناسب أداة القياس المراد تصميمها. وهناك جانب آخر جدير بالذكر في هذا المجال يُشير اليه (عوده، 2002)، هو أن العديد من الباحثين غير المتخصصين في موضوع القياس والتقويم يهمهم إحتساب معامل ثبات الأداة المستخدمة في بحوثهم دون الاهتمام بما تعنيه قيمة ذلك الثبات، وتبرير نوع الثبات المحسوب في ضوء الغرض من المقياس ومصادر الأخطاء المحتملة في القياس وتبرير أهمية النتائج في ضوء خصائص الاختيار ومن بينها قيمة معامل الثبات، فلو كان لدينا اختيارين لهما نفس الهدف وحَسب معامل ثبات كُلٍ منهما بنفس الأسلوب فإن الاختبار الأفضل هو الذي يتميز بأعلى معامل ثبات.

الحكم على معامل الثبات:

للحكم على القيمة المقبولة لمعامل ثبات الأداة علينا أن نتذكر أن معاملات الثبات هي في الحقيقة معاملات إرتباط تتراوح في العادة بين (1) موجب و(1) سالب مروراً بالصفر الذي يعني عدم وجود علاقه، ويؤكد المتخصصون بعدم وجود قاعدة تعمم على جميع أدوات القياس لايجاد تلك القيمة، بل انها تحدد في ضوء عوامل عديدة منها الغرض من الاختبار والأداة، حيث يُشير كل من (Mehrrens and Lehman, 1991) الى أن معاملات ثبات الاختبارات التحصيلية المقننة يجب أن لا تقل عن (0.85) بينما يمكن أن تقل عن تلك القيمة في اختبارات الشخصية.

أما بالنسبة للاختبارات التي يعدها المعلم لطلابه فمعظم معاملات ثباتها تتراوح بين (0.20-0.40) بينما تصل نسبة قليلة منها الى مستوى (0.60).

خصائص الثبات:

أ. الثبات في القياس النفسي يوصف بأنه (نسبي Relative) أي إنه يختلف باختلاف الظروف المحيطة بتطبيقه، فإذا كانت متشابهة تماماً زاد الثبات، وأن اختلفت تؤدي إلى إنخفاضه.

ب. إن قيمة معامل الثبات تمتد بين (صفر، +1)، وكلما كان قريباً من الواحد، يعني أن الأداة جيدة.

جـ. عند تكرار تطبيق الاختبار لإيجاد معامل الثبات قد يتعرض القياس إلى نوعٍ من الخطأ يسمى (خطأ القياس) ينشأ عن مصادر متنوعة يعود بعضها للاختبار نفسه، وبعضها الآخر لظروف تطبيق وتصحيح الاختبار، بينما يعود النوع الثالث من الأخطاء إلى المفحوص الذي يطبّق عليه الاختبار مما يؤدي إلى نتائج غير دقيقة، لذلك يمكن القول أنه كلما زاد مقدار الخطأ قلّ معامل الثبات، وكلما قلَّ الخطأ زاد معامل الثبات.

د. كلما زاد عدد أفراد المجموعة التي يطبق عليها الاختبار مرتين، أو زاد عدد الفقرات (الاسئلة) زادت قيمة تباينهُ، وكلما زاد التباين يزداد معامل الثبات.

العوامل المؤثرة في ثبات الأداة:

أشرنا في الصفحات السابقة أن لأسلوب ايجاد ثبات الاختبار أو أداة القياس أهمية في تحديد مستوى ثبات الاختبار، مما يعني أن لتلك الأساليب تأثيراً ما على تحديد قيمة معامل الثبات، ولكن الأسلوب المستخدم ليس هو الوحيد في هذا الأمر بل ان هناك عوامل عديدة لها دور مهم منها ما يتعلق بالاختبار نفسه مثل صياغة فقراته وطوله وتصميمه، والأفراد الذين يطبق عليهم الاختبار، ثم عوامل تتعلق بالظروف المحيطة بتطبيق ذلك الاختبار . وفيما يلي عرض لتلك العوامل وتأثيراتها على معامل الثبات.

(1) طول الاختبار Test Length

بشكل عام تؤدي الزيادة في طول الاختبار (زيادة عدد فقراته) الارتفاع قيمة معامل ثباته، وذلك لأن الاخطاء العشوائية الموجبة والسالبة التي تصاحب تطبيقه تؤدي إلى استبعاد تلك الاخطاء مما يجعل درجات الطلاب الملاحظه (الناشئة عن عملية تصحيح أوراق اجاباتهم تقترب كثيراً من درجاتهم الحقيقية فيصبح الاختبار أكثر دقةً في قياسه للقدرة المطلوبة. لذلك يمكن القول أن الاختبارات القصيرة تعطي معاملات ثبات منخفضة، وهنا يمكن التوصل الى حقيقة إحصائية تُشير الى أن العلاقة بين طول الاختبار ومعامل ثباته علاقة طردية بمعنى أنه إذا زاد عدد وحداته ارتفع معامل ثباته، ويمكن تمثيل ذلك في المعادلة الآتية التي هي معادلة سبيرمان - براون في صورتها الاولى:

$$\text{ر}_{11} = \frac{\text{ن ر}_{\frac{1}{2}.\frac{1}{2}}}{1 + (\text{ن} - 1)\text{ر}_{\frac{1}{2}.\frac{1}{2}}}$$

حيث أن:

ر_{11}= معامل ثبات الاختبار بعد زيادة عدد فقراته.

ر = معامل ثبات الاختبار قبل زيادة عدد فقراته.

ن = نسبة عدد فقرات الاختبار بعد الزيادة إلى عدد فقرات الاختبار قبل الزيادة.

مثال:

فلو فرضنا أن لدينا اختبار عدد فقراته (20) فقره، وكان معامل ثباته (0.60)، وأردنا مضاعفة طوله ثلاث مرات عن طوله الحالي أي أن تكون عدد فقراته (60) فقرة، فإن معامل الثبات الجديد سيصبح :

$$ر_{1-1} = \frac{0.60 \times 3}{0.60\ (1-3) + 1} = 0.818$$

ونلاحظ قيمة معامل الثبات زادت عندما زادت عدد فقراته.

ويمكن الاستفادة من نفس المعادلة السابقة عندما نريد معرفة قيمة (ن) أي معرفة النسبة التي يجب تحقيقها (أي مستوى زيادة عدد الفقرات) للوصول إلى درجة معينة من الثبات، والمثال التالي يوضح ذلك.

مثال:

لنفرض أن لدينا اختبار عدد فقراته (50 فقرة) ومعامل ثباتهُ (0.70) ونريد زيادة معامل الثبات الى (0.90) فكم يجب أن يكون عدد فقراته؟.

وبتطبيق المعادلة السابقة نصل إلى الجواب المطلوب بعد الأخذ بالاعتبار تعويض قيمة الثبات المطلوبة بالعدد (0.90) إذن:

$$0.90 = \frac{0.70 \times ن}{0.70\ (1 - ن) + 1}$$

ومن خلال نقل القيم الى مواقع مختلفة تجد أن:

$$0.70 ن= 0.90 + (0.9 ن - 0.9)\ 0.7$$

$$0.70 ن = 0.90 + 0.63 ن- 0.63$$

$$0.70 ن = ن- 0.63 ن= 0.270$$

$$0.70 ن = 0.270$$

ن = 4 تقريباً.

أي إذا أردنا الحصول على معامل ثبات مقداره (0.90) فعلينا زيادة عدد الفقرات الاختبار الحالية إلى أربعة أمثالها أي إلى (200) فقره.

(2) تجانس المجموعة Croup Homogeneity

إذا تم تثبيت العوامل الاخرى في عملية تقدير الثبات، فان مجموعة الدرجات التي يكون تباينها أعلى هي التي ينتج عنها معامل ثبات أكبر. ويمكن توضيح العلاقة بين التباين والثبات في المعادلة التالية:

$$r_{x.x} = 1 - \frac{S^2y}{S^2x}(1 - r_{y.y})$$

أي أن معامل الثبات $ر_{ص.ص} = \frac{\text{تباين المجموعة الأولى}}{\text{تباين المجموعة الثانية}} (1 - ر_{س.س})$

أو $ر_{ص.ص} = 1 - \frac{ع^2_{س}}{ع^2_{ص}} (1 - ر_{س.س})$

مثال:

فلو طبعنا اختباراً ما على مجموعة من الافراد (س) ووجدنا أنه يساوي (0.60) وان تباين درجات تلك المجموعة = 16. وطبقنا نفس الاختبار على مجموعة اخرى (2)، ووجدنا أن تباين درجات تلك المجموعة = 25 فإن معامل ثبات الاختبار المطبق على المجموعة الثانية (ص)=

$$ر_{ص.ص} = 1 - \frac{16}{25}(1 - 0.60)$$

$$ر_{ص.ص} = 1 - 0.26$$

$ر_{ص.ص} = 0.74$ ويلاحظ ان معامل ثبات الاختبار ارتفع عندما طبق على مجموعة تباينها أعلى من تباين المجموعة الاولى.

إن العلاقة الطردية بين تباين درجات المجموعة ومعامل ثبات الاختبار تحدث بسبب قدرة الاختبار على التمييز بين الفروق الفردية لأفراد المجموعة التي تؤدي الى زيادة درجاتهم بما يؤدي الى زيادة معامل ثبات الاختبار (العاني، 2003).

(3) مستوى صعوبة الفقرات Items Difficulty

يتأثر معامل ثبات الاختبار بمستوى صعوبة فقراته، فإذا ما علمنا من خلال ما ذُكرَ أعلاه أن ثبات الاختبار يعتمد على مستوى تباينه، فان فقرات الاختبار إن كانت سهلة جداً فإن جميع الطلاب أو أغلبهم سيجيبون عليها إجابة صحيحة، وإن كانت صعبة جداً فإن أغلبهم سيجيب عليها إجابة خاطئة وبذلك ينخفض تباين درجات الاختبار مما يؤدي إلى إنخفاض معامل الثبات، لذلك يُفترض بمصمم الاختبار أن يحافظ على المستوى المقبول لصعوبة الاختبار وهو المستوى المقترب من (0.50).

(4) الزمن المخصص للأختبار Test Time

يتأثر ثبات الاختبار بالمدة المخصصة للإجابة عليه، وقد وجد الباحثون أن معامل الثبات يزداد لزيادة الوقت المخصص له حتى يصل إلى الحد المناسب للاجابة وعندها يرتفع معامل الثبات إلى أعلى مستوى، لكنه يبدأ بالانخفاض عندما يزداد الوقت عن الحد المناسب. وقد لوحظ أيضا أن الاختبارات المحدد زمن الاجابة عليها تميل معاملات ثباتها للأرتفاع أكثر من الاختبارات ذات الزمن المفتوح الذي يفوق الوقت الذي يحتاجه الفرد ذو القدرة المتوسطة للاجابة عليه.

وبالاضافة إلى ما ذُكر أعلاه فهناك عوامل أخرى منها على سبيل المثال لا الحصر التخمين الذي يستخدمه الطالب أثناء الاجابة الذي قد يؤدي إلى إنخفاض معامل الثبات، وكذلك حالة الطالب الجسمية والنفسية أثناء إدائه الاختبار، وموضوعية تصحيح الاختبار وظروف اجراء وإدارة الاختبار المتعلقة بالظروف المحيطة بالطالب أثناء الاختبار كإسلوب تعامل المراقبين على الاختبار مع الطلبة والضوضاء ودرجة الحرارة والاضاءة والتهوية وغير ذلك.

العلاقة بين الصدق والثبات

من خلال ما تمَّ تقديمه عن مفهومي الصدق والثبات لا بد للقارئ أن

يستنتج إنَّ أبسط أنواع العلاقة بين المفهومين إنهما يمثلان خاصيتين تتعلقان بمفهوم واحد هو المقياس أو الاختبار لا سيما كفاءته في إداء المهمة التي يستخدم لتحقيقها. فاذا ما علمنا أن الثبات يعني إستقرار درجات المجموعة التي تجيب عن الاختبار عندما تتغير الظروف الخارجية فان ذلك يعني أن الثبات يختص بالاختبار ودرجاته. أما مفهوم الصدق فانه يتجاوز الاختيار ودرجاته الى درجات إختبار آخر هو المحك الذي نستخدمه من أجل المقارنة والحكم من خلال استخدام معاملات الارتباط بين درجات الاختبار ودرجات المحك. ونلاحظ أننا في كلا عمليتي الثبات والصدق نعتمد اعتماداً كليا على درجات المجموعة أي على نتائج الاختبار نفسه فهي العامل المشترك في ايجاد كلا المعاملين. من ناحية أخرى نلاحظ أن الصدق يعني تمثيل الاختبار للمسة المقاسة تمثيلاً حقيقياً، وفي الثبات نريد التأكد من أن هذا التمثيل يبقى مستقراً عبر الزمن أو عبر أساليب أخرى، ومن هنا يمكن الاستنتاج أن الاختبار الصادق يبقى صادقاً عبر الزمن أي سيكون ثابتاً قليل التغير في تمثيله للصفة المقاسة إن كان هناك نوع من التغير، لكن ذلك لا يشترط للاختبار الثابت أن يكون صادقاً دائماً، فعلى سبيل المثال لو قمنا بقياس أطوال مجموعة من الاطفال وحصلنا على درجات تمثل تلك الاطوال لمرتين، وقمنا بإدخالها في الحاسب الآلي على إنها أوزان لمجموعة من الأطفال قيست مرتين فسوف نحصل على معامل ثبات لتلك الدرجات ذو مستوى عال، ولكن هل إن ذلك يعني أن هذه القياسات صادقة؟ أي هل هي تمثل الأوزانَ حقا؟ الجواب كلا، ومن هنا نستنتج أن القياسات الثابتة لا يشترط أن تكون صادقة. بينما القياسات الصادقة تكون ثابتة على الأغلب. من ناحية أخرى يرى بعض الباحثين أن العلاقة بين الصدق والثبات علاقة واضحة في الجانب الاحصائي بشكل أكبر، حيث يؤثر معامل الثبات على معامل الصدق، يؤكد ذلك ما يطلق عليه الاحصائيون "أقصى معامل صدق" لا سيما في صدق المحك حيث يساوي هذا المعامل الجذر التربيعي لحاصل ضرب معامل ثبات الاختبار الجديد بثبات اختبار المحك، أي أن

أقصى معامل الصدق= $\sqrt{\text{معامل ثبات الاختبار الجديد} \times \text{معامل ثبات المحك}}$

(العاني، 2003)

ثالثاً - الموضوعية Objectivity

تعتبر الموضوعية من الخصائص المهمة في أداة القياس، ولمصطلح الموضوعية عدة معان منها إنها عكس الذاتية، أي أنها العملية التي لا تتأثر بالعوامل الذاتية للفرد. وفي القياس يقصد بالموضوعية الإتفاق في الأحكام، أي ان ما يحكم عليه (س) من الناس يتفق مع الحكم الذي يصدره الآخرون على نفس الظاهرة. وبشكل أكثر اجرائياً إنها تعني الاتفاق بين المحكمين في الحكم الذي يصدرونه في عملية التقويم- وبالنسبة للاختبارات نقصد الموضوعية إتفاق مجموعة المصححين على إعطاء نفس الدرجة لورقة إجابة الطالب مهما إختلف عددهم. إن المقياس أو الاختبار يكون موضوعياً عندما لا يوجد إختلاف لذلك يمكن الاستنتاج بأن الموضوعية تحقق مبدأ العدالة في تحديد الدرجة التي يستحقها الطالب دون زيادة أو نقصان، أو تدخل عوامل ذاتية أو مزاجية للقائم بعملية التصحيح. أو تدخل عوامل ذاتية أو مزاجية للقائم بعملية التصحيح. ويمكن تحقيق الموضوعية من خلال صياغة فقرات الاختبار بحيث تكون إجابة الطالب عليها إما صحيحة تماماً أو خاطئة تماماً وهذا ما نراه في الاختبارات الموضوعية Objective Tests مثل اختبارات الصواب والخطأ، والاختبار من متعدد واختبار المطابقة أو المزاوجة. ولا يتحدد مفهوم الموضوعية في الامتحانات التحريرية فقط بل يجب أن يمتد الى جميع أدوات القياس الاخرى لكي تنتج عنها أحكاماً أو تقويمات موضوعية ولذلك يركز العاملون في مجال القياس التربوي على ضرورة صياغة فقرات تلك الأدوات وفقاً للقواعد العلمية لصياغتها. ففي مجال الاختبارات التحصيلية يجب التقليل من أسئلة المقال لأنها تتأثر كثيراً بالعوامل الذاتية، كما يمكن إستبدالها بأسئلة ذات الاجابات القصيرة أو المحددة.

رابعاً - الشمول Comprehensiveness

يقصد بالأداة الشاملة تلك التي تمتد فقراتها لقياس جميع أبعاد الخاصية المراد قياسها. فلو كان لدينا مقياس لقياس ذكاء الأطفال في عمر ست سنوات فان فقرات ذلك المقياس يجب أن تمثل كافة أبعاد ذكاء الأطفال في عمر الست سنوات.

وعندما نريد تصميم اختبار تحصيلي لقياس تحصيل الطلاب في مادة الرياضيات في الصف السادس الابتدائي فإن فقرات ذلك المقياس يجب أن تغطي جميع أبعاد محتوى المادة الدراسية (مادة الرياضيات) التي درست للطالب في فترة ما، اضافة الى أن تلك الفقرات يجب أن تمثل الانماط السلوكية المراد تقويمها لدى أولئكَ الطلاب، وبذلك يمكن القول أن هذا الاختبار شمل محتوى المادة والأهداف السلوكية المراد تقويمها من خلال ذلك الاختبار، ونظراً لأن المادة الدراسية غالباً ما تكون محددة المعالم من خلال الكتاب المدرسي المطبوع فان عملية تحديد المحتوى سوف لن تكون صعبة على مصمم الاختبار التحصيلي، لكن الأمر أكثر صعوبة في تصميم الاختبارات النفسية والاجتماعية كإختبارات الميول والقدرات والاتجاهات وغير ذلك بسبب عدم وجود محتوى محدد ومطبوع كما في الكتاب المدرسي، يزيد الأمر صعوبة عدم وجود اتفاق بين العلماء والباحثين على تعريفات واطار فكري ينتمي اليه وينظر من خلاله مما يعقد الموضوع على الباحث الذي يريد تصميم اختبار أو مقياس نفسي جديد، لذلك عليه أن يوائم بين النظريات المختلفة أو يلتزم بإطار فكري أو نظري معين يحدد من خلاله مفاهيمه وحدود محتوى المقياس الذي يريد تطويره، وكثيراً ما يلجأ بعض الباحثين وطلبة الدراسات العليا للتخلص من هذا المأزق الى أساليب عدة منها الاستعانة بالمقاييس الشائعة المعروفة للاقتباس من محتوياتها مما يُزيد التشابه بين المقاييس المختلفة ويضعف عامل الحداثة والابداع، أو الاعتماد على آراء مجموعة من المحكمين لاعطاء الباحث ملاحظَاتهم عن محتوى المقياس الذي يقوم بتصميمه وفي كثير من الأحيان تتم عملية التحكيم بالرتابة وعدم الدقة وهذا ما نلاحظه على تصميم الكثير من مقاييس بحوث طلبة الدراسات العليا التي ينقصها الكثير

من الدقة والشمول وتمثيل السمة المراد قياسها بشكلٍ موضوعي ودقيق. إن خاصية الشمول من الخصائص المهمة جداً في عملية بناء الاختبارات والمقاييس لأنها ترتبط إرتباطاً وثيقاً بعملية صدق المقياس، فالمقياس الذي لا يشمل جميع أبعاد السمة المقاسة سيهمل جوانب هامة من تلك السمة أي عدم تمثيلها فيه مما يجعل الأختبار غير صادق في قياس ما وضع لقياسه. ونظراً لأن المحور الأساسي الذي نركز عليه في هذا الفصل هو الاختبارات التحصيلية المستخدمة لقياس التحصيل المدرسي، فان الشيء الذي على المعلم أو الباحث الذي يريد تصميم الاختبارات التحصيلية الاهتمام به لتحقيق خاصية الشمول هو وضع ما يسمى بـ (خطة الاختبار Test Plan) أو (جدول المواصفات Table of Spesifications).

وفيما يلي توضيح للهدف من هذا الجدول وأسلوب تصميمه.

جدول المواصفات Table of Specifications

مما لا شك فيه إن أي عملٍ يكون أكثر تنظيماً ومعاليةً عندما يسبقه نوع من التخطيط الذي يتضمن وضع حسابات منظمة ودقيقة لتنفيذ ذلك العمل.

وهذا ينطبق على عملية تصميم أو بناء الاختبارات بكافة أنواعها لا سيما الاختبارات التحصيلية. فمن الآليات المهمة في بناء الاختبار هو وضع خطة أو مخطط لذلك الاختبار قبل البدء بكتابته من خلال ما يسمى بـ (جدول المواصفات) الذي يطلق عليه عادةً Blue Print أو المطبوع الأزرق، هذا المصطلح أُستعير من مخطط البناء الأزرق الذي يضعه المتخصصون في تصميم المنازل والعمارات لأجل أن يسير عليه البنائون في تنفيذهم لتلك المنازل أو العمارات. إن هذا المخطط بلا شك سيساعد المعلم أو من يقوم بتصميم الاختبار بأخذ الجوانب الأساسية بنظر الأعتبار من أجل أن يظهر الاختبار بأفضل صورة تتمثل فيه متطلبات وشروط الاختبار الجيد لا سيما الشمول حيث سيتضمن أهم عنصرين من عناصر الشمول وهما:

أ- تغطية محتوى المادة المراد قياس مدى إتقان الطالب لها.

ب- تحديد الأهداف السلوكية المراد قياسها من خلال الأختبار.

إن تصميم جدول المواصفات يتطلب إيجاد نوع من التوازن بين البعدين الأساسيين في الجدول وهما المحتوى والأهداف وفقاً للخطوات الآتية:

1 - تقسيم المحتوى (المادة التعلمية) الى موضوعات أو وحدات وتحديد الوزن النسبي لكل وحده من وحدات المحتوى وفقاً لاحدى الطرق التاليه:

أ- الاعتماد على عدد الصفحات التي يتضمنها الكتاب المخصصة للوحدة التي سيغطيها الاختبار من الكتاب. فاذا كان عدد صفحات الوحدة الاولى- مثلاً- هو 15 صفحة وكان عدد الصفحات المقررة لعموم مادة الاختبار 60 صفحه فان نسبة محتوى هذه الوحدة = $\frac{15}{60} \times 100 = 0.25$
وهكذا بالنسبة للموضوعات أو الوحدات الباقية.

ب- عدد الأهداف المحددة في كتاب (مرشد أو دليل المعلم) المخصصة لتلك الوحدة وايجاد نسبتها الى العدد الكلي لأهداف الوحدات التي سيغطيها الاختبار.

فإذا كان عدد أهداف الوحدة الأولى مثلاً (12) ومجموع الأهداف الكلية للوحدات هو (80)، فان نسبة أهداف الوحدة $\frac{12}{80} \times 100 = 0.15$ وهكذا.

ج- وهناك من يحسب النسبة وفقاً لعدد الحصص التي يخصصها المعلم لتدريس وحدة معينة وقسمته على العدد الكلي للحصص المخصصة للوحدات التي يشملها الاختبار. فلو كان عدد الحصص المخصصة لتدريس وحدة معينة هو (4) حصص وان المجموع الكلي للحصص المخصصة للوحدات التي يشملها الاختبار هو (36) حصة فان النسبه $= \frac{4}{36} \times 100 = 0.11$

2 - أما بالنسبة لتحديد الوزن النسبي الخاص بالأهداف التعليمية (نواتج التعليم) ولكل هدف على حده، فان ذلك يعود الى رأي المعلم الذي يقوم بتدريس المادة الدراسية على اساس أنه خبر المادة من جهة، ومستويات طلابه وأهمية الهدف المراد تحقيقه أو قياسه فهو أقدر من غيره

على تحديد تلك النسبة، أما الأسلوب الآخر الممكن اتباعه في تحديد الأوزان النسبية لنواتج التعلم فهو الاعتماد على بعض المتخصصين من اساتذة المادة والمشرفين التربويين وأساتذة الجامعات حيث يتم عرض الموضوعات عليهم وأخذ آرائهم وإتفاقهم على الأوزان النسبية التي يحددونها لكل وحدة أو موضوع دراسي.

مثال:

أراد أحد المعلمين أن يصمم إختباراً تحصيلياً لطلاب الصف السادس في مادة تدريسية معينة يتألف من (50) فقره من النوع الموضوعي. يشمل هذا الاختبار ثلاث وحدات تغطي (40) صفحة من الكتاب المقرر، وثلاثة أهداف سلوكية من تصنيف "بلوم" للأهداف السلوكية هي:

المعرفة، الفهم، والتطبيق - ولأجل أن يضع جدول مواصفات لهذا الاختبار إتبع الخطوات التالية:

1- تحديد الوزن النسبي لمحتوى المادة من خلال معرفة عدد صفحات كل من الوحدات الثلاثة وكانت كالاتي:

- عدد صفحات الوحدة الأولى هي (20) صفحة وبذلك يكون

الوزن النسبي لهذه الوحدة = $\frac{20}{40} \times 100 = 0.50$

- عدد صفحات الوحدة الثانيه هو (12) صفحه، وبذلك يكون

الوزن النسبي للوحدة = $\frac{12}{40} \times 100 = 0.50$

- عدد صفحات الوحدة الثالثة هو (8) صفحات، وبذلك يكون

الوزن النسبي للوحدة = $\frac{8}{40} \times 100 = 0.50$

2- تحديد الوزن النسبي لكل هدف من الأهداف السلوكية التي يريد قياسها لدى الطلبة. وبناءً على خبرته الشخصية ومعرفته لمستوى طلابه وطبيعة الماده الدراسيه، فقد وضع الاوزان النسبية للاهداف الثلاثة كالآتي:

الوزن النسبي للمعرفة = 0.60

الوزن النسبي للفهم والاستبيان = 0.15

الوزن النسبي للتطبيق = 0.25

وبعد أن تمَّ تحديد العدد الكلي لفقرات الاختبار بـ (50) فقره وتحديد الأوزان لكلٍ من المحتوى والاهداف يبقى أن يحدد عدد الفقرات الخاصة لكل خلية من خلايا جدول المواصفات الذي سيبلغ عددها تسع خلايا والتي سترمز لها (ن) من خلال تطبيق المعادلة التالية:

ن - الوزن النسبي للوحدة × الوزن النسبي للهدف × العدد الكلي لفقرات الاختبار

عدد فقرات الخلية الاولى = 0.50 × 0.60 × 50 = 15 فقره وهي التي ستصمم من الوحدة الاولى وتقيس قدرة الطالب على المعرفه.

عدد فقرات الخلية الثانية = 0.50 × 0.15 × 50 = 3.75 وهي التي تصمم من الوحدة الاولى وتقيس الفهم لدى الطلبة.

عدد فقرات الخلية الثالثة = 0.50 × 0.25 × 50 = 6.25 فقرة تصمم من الوحدة الاولى لقياس القدرة على التطبيق لدى الطلبة.

ويمكن إتباع نفس الاسلوب لايجاد عدد فقرات الخلايا الباقية من الجدول حيث سيكون العدد المقرر للخلية الرابعة $ن_4$ = 9 فقرات.

$ن_5$ = 2.25 فقره، $ن_6$ = 3.75 فقره، $ن_7$ = 6 فقرات، $ن_8$ = 1.5 فقره، $ن_9$ = 2.5 فقره.

ولو قمنا بجمع أعداد الفقرات التي تتضمنها خلايا الجدول أفقياً وعمودياً لوجدنا أنها تساوي 50، وهو العدد الكلي لفقرات الاختبار.

ويمكن توضيح الخطوات السابقة في الجدول الآتي:

المجموع	التطبيق (0.25)	الفهم (0.15)	المعرفة (0.60)	المحتوى/الأهداف
25	(3) 6.25	(2) 3.75	(1) 15	الوحدة الأولى (0.50)
15	(6) 3.75	(5) 2.25	(4) 9	الوحدة الثانية (0.30)
10	(9) 2.5	(8) 1.5	(7) 6	الوحدة الثالثة (0.20)
50	12.5	7.5	30	المجموع 100%

ويلاحظ على الجدول السابق إحتوائه على الارقام التي تدل على عدد الفقرات التي ستخصص لكل خلية بعضها يتكون من كسور عشرية، مما يمكن أن يُثير تساؤلاً عن كيفية تصميم أجزاء من الأسئلة (الفقرات) إذا أردنا تحقيق محتوى الجدول تحقيقاً كاملاً عند تصميم فقرات الاختيار وتلافياً لذلك يمكن للباحث أو المعلم أن يعمل نوعاً من التوازن بين الكسور العشرية في الخلايا من خلال رفع كسر عشري من خلية واضافتة لكسر عشري آخر في خلية أخرى ويفضل أن يكون ذلك ضمن الوحدة الواحده، وبذلك يضمن وجود أرقام صحيحه تمثل أعداد الفقرات التي يجب ان تصمم لكل خلية بحيث تمثل محتوى من وحدة معينة لتقيس هدفاً محدداً. ويمكن توضيح ذلك في إعادة صياغة الجدول السابق بالصورة الجديدة المدرجة في أدناه:

المجموع	التطبيق (0.25)	الفهم (0.15)	المعرفة (0.60)	المحتوى/الأهداف
25	6	4	15	الوحدة الأولى (0.50)
15	4	2	9	الوحدة الثانية (0.30)
10	2	2	6	الوحدة الثالثة (0.20)
50	12	8	30	المجموع 100%

إن جدول المواصفات سيضع أمام الباحث او المعلم خطة مدروسة لتحقيق شمولية الاختبار وبالتالي صدقه في قياس ما وضع لقياسه وسيكون دليلاً ومرشداً له يسير في ضوئه عند تصميم اسئلة الاختبار، وننصح معلمي المدارس بكافة مراحلها بضرورة تصميم مثل هذا الجدول منعاً للوقوع في أخطاء كثيرة تضعف من الثقة بنتائج الاختبار، سواء في صدقه أو ثباته، لأن عدم تصميم مثل هذا الجدول يدفع المعلم بشكل خاص الى الاعتماد على العامل الذاتي، أو عامل الصدفة في تفضيل هذا الموضوع على ذاك من الموضوعات التي درسّها لطلابه وبالتالي الوقوع بالعديد من الأخطاء في عملية التصميم.

لقد تم في الصفحات السابقة تقديم عرض لنموذج جدول مواصفات عام أي غير تفصيلي من حيث محتوى المادة او الاهداف التدريسية، لكن ذلك لا يمنع المعلم ليكون عمله أكثر دقة كأن يدرج الموضوعات التفصيلية التي تتضمنها وحدات المادة الدراسية، وبعض آليات أو تفصيلات كل هدف يريد تحقيقه بالاختبار المصمم، ونقدم فيما يلي نموذجاً لجدول مواصفات يوضح ذلك ويتعلق بموضوعات تتعلق بمادة القياس والتقويم.

المحتوى/الأهداف	معرفــــة			فهم المبادىء	تطبيق المبادىء	تطبيق المبادىء
	مصطلحات	حقائق	أساليب			
دور الاختبارات في التدريس	4	4	-	2	-	10
أنواع الاختبارات	4	3	2	6	5	20
الاختبارات المحكية والاختبارات المعيارية	4	3	3	-	-	10
تخطيط الاختبار	3	5	5	2	5	20
المجموع	15	15	10	10	10	60

الأهداف: Objectives

يقصد بالاهداف الغايات التي ترمي الوصول اليها أو تحقيقها من خلال عمل معين. وللأهداف فوائد كثيرة ومتنوعة لكنها بشكل خاص تعمل كمرشد أو مُوجه للأعمال التي يزاولها الفرد. ويجمع الكثير من الباحثين على أن الأهداف عندما تكون واضحة ومحددة فانها تدفع الفرد الى اختيار الوسائل المناسبة لتحقيقها وتقلل من فرص إستهلاك الوقت والجهد بل تختزلهما الى أقّل حد ممكن وتؤدي الى التنظيم الذي يوصل للهدف بصورة أسرع من التخبط واللف والدوران اللذان كثيراً ما يؤديا الى عدم تحقيق الأهداف.

لقد ركزت العديد من نظريات علم النفس على ضرورة أن تكون أهداف الفرد واضحة ومحددة وواقعية حتى يسهل عليه تحقيقها أو بلوغها، وهناك من مدارس علم النفس وهي المدرسة الغرضية أو الغائية Purposive School التي تؤكد على أن سلوك الفرد يوجه دائماً نحو أهداف أو أغراض محددة تشبع حاجاته الداخلية وتحفزه للاستمرار في العمل لتحقيق تلك الأهداف.

وفي العملية التعليمية تبرز الحاجة الى وجود أهداف واضحة ومحددة أكثر مما في الأعمال والمهام الأخرى التي يزاولها الفرد، فالمعلم الذي لا يحدد أهدافه بوضوح يتخبط في عملية تدريسه لطلابه ويقلل من فرص تعلمهم بالمستوى المطلوب. ومن هنا أصبح على المعلم أن يحدد أهداف تدريس المادة الدراسية ضمن خطة التدريس اليومية ويبين الغايات التي يريد تحقيقها من تدريس الموضوع. كما أن عليه أن يحدد تلك الأهداف ضمن خطة الاختبار الذي يريد تصميمه ويربطها بشكل محكم بمحتوى المادة الدراسية كما لاحظنا ذلك في تصميم جدول المواصفات، لأن عليه أن يحدد تلك الأهداف ضمن خطة الاختبار الذي يريد تصميمه ويربطها بشكل محكم بمحتوى المادة الدراسية كما لاحظنا ذلك في تصميم جدول المواصفات لأن الأرتباط بين محتوى المادة الدراسية والأهداف المراد قياسها يعتبر ارتباطاً عضوياً يعمل المعلم جاهداً لتحقيقه من خلال عمليتي التدريس والأختبار.

ويتبع في الميدان التربوي مصطلح "الأهداف التعليمية" أي ما نريد تحقيقه من خلال العملية التعليمية ويُعبَّر عنهُ بصوره لفظيه أو بعبارات ترسم صورة المتعلم بعد إنجاز عملية التعليم. وعندما يضاف الى الأهداف مصطلح (السلوكية) لتُقرأ (الأهداف السلوكية) فإنما يقصد بها صفات محددة ومرغوب فيها يجب أن تظهر في سلوك المتعلم ومن الممكن ملاحظتها أو قياسها في نهاية التدريس أو في مرحلة معينة من مراحل ذلك التدريس. (العاني، 2003).

ونظراً لوجود مصطلحات عديدة تضاف الى مصطلح الأهداف في مجال التربية والتعليم نقدم فيما يأتي توضيحاً لمستويات تلك الأهداف لمساعدة الطالب على التمييز بينها وإستخدام كلٍ منها في المكان المناسب.

مستويات الأهداف التربوية:

١- **الأغراض التربوية** (Educaational Purposes) وهي النتائج النهائية المرغوبة من الناحية التربوية . وتتسم بأنها شديدة العمومية وتستخدم للتعبير عن البرامج التربوية بشكل عام. مثال: إعداد المواطن الصالح.

2- **الغايات التربوية** Educational Aims وهي عبارات تصف نتاجات حياتية متوقعة ومبنية على مخطط قيمي مشتق من الفلسفة التربوية السائدة في المجتمع. وتمثل أهداف بعيدة المدى تتعدى مستوى الموقف الصفي. ويتم إعداد قائمة الغايات بواسطة: مجالس التربية والمجتمع والمختصين في النظام التربوي.

3- **الأهداف التربوية العامة** وتشمل الأهداف المنشودة من تدريس مادة دراسية ما. مثال لهدف عام من تدريس التربية الإسلامية: تنمية الاتجاهات والأخلاقيات التي يشجع عليها الدين الإسلامي.

4- **الأهداف التعليمية** وتسمى أيضا (السلوكية) وتمثل العبارات التي يسعى المعلم والتلاميذ إلى تحقيقها عند تدريس درس معين وتعرف اجرائياً بأنها التغيرات الايجابية المطلوب احداثها في سلوك الطالب. وتستمد الأهداف السلوكية عامة من الأهداف العامة للمنهج.

أهمية الأهداف التعليمية أو السلوكية:

1- تساعد المعلم في التخطيط للدرس.

2- تنظيم العملية التعليمية. وذلك لأن المعلم والطالب يجب إن يكونا على بينة بالسلوك او النتاج النهائي والمطلوب تحقيقه من التدريس.

3- تساعد المعلم على وضع آلية التقويم.

4- تجزئة وتحليل محتوى المادة إلى أقسام صغيرة يمكن توضيحها وتفسيرها للطلاب.

5- تساعد المعلم على اختيار طريقة التدريس المناسبة والوسيلة التعليمية الملائمة.

6- تساعد على تطبيق الأهداف العامة للمنهج.

وتتم وفقاً للخطوات الآتية:

1- تحديد المعارف او المهارات المراد إكسابها للطالب، أو تحديد المحتوى.

2- تحديد السلوك الدقيق.

3- تحديد النتيجة المرغوب فيها وهذه تقتضي أن يحدد المعلم درجة الدقة أو احتمالية الخطأ في السلوك النهائي.

4- تحديد الشروط التي توضح طريقة إثبات السلوك أو البرهنة عليه. وهنا يتم تحديد الوسائل التعليمية التي يمكن استخدامها لتحقيق الهدف.

الهدف التعليمي (السلوكي الجيد):

1- يجب أن يصف ما يقوم به المتعلم خلال الحصة أو عند الانتهاء منها.

2- يجب أن يصاغ بشكل يجعله قابلاً للقياس.

3- يجب أن يشتمل الهدف على ثلاثة عناصر مهمة وهي:

أ - السلوك (يذكر، يفسر، يقارن).

ب- (الظروف أو الشرط (بالرجوع الى الكتاب).

ج- (المعيار مثلاً (80%. بدقة تامة. بدون أخطاء).

مثال لهدف سلوكي:

أن يتلو الطالب سورة الأخلاص وفقاً لأحكام التجويد والنطق السليم وبنسبة صواب لا تقل عن 90%.

تصنيفات الأهداف السلوكية: حدد بلوم (Bloom, 1956)، في تصنيفه للأهداف السلوكية ثلاثة مستويات وهي:

أولاً: الأهداف في المستوى المعرفي: Cogntive damain وتصنف أهداف هذا المجال الى ستة مستويات وهي:-

- التذكر او المعرفة: أن يعدد الطالب الكواكب القريبة من الشمس، بعد الاطلاع على الوسيلة التعلمية، وبدون أخطاء.
- الفهم والاستيعاب: أن يفسر الطالب ظاهرة المد والجزر، بعد مشاهدة الفلم التعليمي، وبنسبة صواب لا تقل عن 90%.
- التطبيق: أن يستخرج معامل الارتباط بمعلومية المعادلة وبدقة.
- التحليل: يقارن بين غزوتي أحد والأحزاب من حيث الأسباب والتطورات والنتائج وبنسبة صواب لا تقل عن 80%.
- التركيب: يقترح خطة لمعالجة مشكلة تملح الآبار في منطقة جغرافية معينة في ثلاث صفحات.
- التقويم: يحكم على الدور الذي لعبه الصحابي خالد بن الوليد في حروب الردة بعد مناقشة دوره.

الأهداف في المستوى الانفعالي Affective Damain

وقام بتصنيفها التربوي (كراثول) وتنقسم الى خمسة مستويات. وفيما يلي مثال لكل مستوى من المستويات الخمسة:

- الاستقبال أن يهتم الطالب بقضية اضطهاد الشعب الفلسطيني من قبل قوات الاحتلال الاسرائيلي.
- الاستجابة: ان يتطوع للمساهمة في جمع التبرعات للفقراء في المجتمع إذا ما طلب منه.

- التقييم (مرحلة تكوين الاتجاهات والمعتقدات): يناقش الآثار الاجتماعية والاقتصادية للعولمة على المجتمع المحلي إذا ما طرح الموضوع.
- (تدخل القيمة ضمن النسق القيمي للفرد): يخطط لحل مشاكل الفقر في المجتمع باستخدام أموال الزكاة.
- تشكيل الذات (يتكون نظام قيمي لدى الفرد يتحكم في أسلوب حياته).

أن يحترم الطالب والديه ويرعاهما في ضوء فهمه لتعاليم القرآن.

الأهداف في المستوى المهاري الحركي (النفسحركي) Psycho-motor Domai

وحددها بعض التربويين مثل جرونلند، هارو، سمبسون. وتنقسم أهداف هذا المستوى حسب تصنيف سمبسون إلى:

- الإدراك الحسي.
- الميل والاستعداد.
- الاستجابة الموجهة.
- الآلية والتعود.
- الاستجابة الظاهرية المعقدة.
- التكيف أو التعديل.
- الابداع.

خامساً - التقنين Standardization

يعتبر التقنين من المستلزمات التي يتطلبها الاختبار الجيد. ويلاحظ أن الكثير من المعلمين يستخدمون أساليب تقليدية عندما يقومون بمقارنة الدرجة التي يحصل عليها طالب معين مع درجة طالب آخر، والاستنتاج مثلاً ان درجة أحمد في مادة التاريخ (86) أفضل من درجة حسين الذي حصل على 80 ، بينما هذه المقارنة غير مقبولة تربوياً لأن درجة أحمد وحسين تسمى بالدرجة الخام Raw Score

والدرجة الخام ليس لها معنى، ولا يمكن استخدامها للمقارنة والحكم واتخاذ القرار وما شابه ذلك، إلا بمقارنتها بمتوسط المجموعة التي طبق عليها الاختبار، وهذا المتوسط يمثل الاداء الطبيعي او العادي.

إن مفهوم القياس يتضمن جانبين، الاول هو اعطاء الارقام للخصائص، والثاني وجود نظام او قواعد للارقام التي تم إعطاؤها. وفي حالة الدرجة الخام، فانها تمثل الجانب الاول فقط وهو "اعطاءها الارقام"، وليس نظام او قواعد الارقام، وعملية وضع القواعد لتفسير الأرقام هي التي تمثل عملية التقنين. فما هو التقنين؟ وكيف يمكن التحقق منه؟

يقصد بالتقنين توحيد الاجراءات وتشابهها، وفي أدبيات القياس له تعريفات عديدة. فقد عرفه (ماكماهون) بأنه عملية تهدف إلى جعل الاختبار منسجماً عبر تطبيقه من مرحلة إلى أخرى، أي أنه يتحول إلى اختبار مقنن (Mcmahan,1974)، وتعرفة انستازي واربينا على أنه وضع شروط موحدة لتطبيق الاختبار على جميع الافراد، كما يتضمن طريقة موحدة في تقييم الاستجابات .(Anastasi & Urobian, 1997)

ويرى ايبل (Ebel,1972) مواصفات الاختبار المقنن في ثلاث نقاط:

1- يوضع بعناية فائقة من خلال خبرة عميقة، ويمر بتجارب استطلاعية وتحليل واعادة.
2- تكون له تعليمات واضحة وموحدة (مقننة) للجميع.
3- تكون له جداول للمعاير طبقت على عينة معروفة من الطلبة.

وهكذا فإن عملية التقنين هي إحدى الخواص المهمة للاختبار التحصيلي عندما يراد تطبيقه على عينة تتجاوز حدود المدرسة إلى مجموعة من المدارس، أو ربما على مستوى المنطقة التعليمية، ومن أبرز الأمثلة على الاختبارات المقننة على المستوى المحلي (اختبار الثانوية العامة) التي تعده في الغالب لجنة متخصصة في وزارة التربية والتعليم..

وتعد المعايير Norms الخاصية المشتركة بين الاختبار المقنن (كما وردت لدى ايبل) واختبارات المعلم. وهي خطوة لا بد من القيام بها بعد التأكد من صدق الاختبار وثابته.

والمعايير هي موازين مستوى اداء مجموعة من الطلبة في اختبار ما، تستخدم للحكم على مستوى اداء أي طالب بالنسبة لأداء المجموعة التي ينتمي اليها.

وطبقاً للجماعة المرجعية (عينة الطلاب الذين طبق عليهم الاختبار) تقسم المعايير الى اربعة اقسام هي:

1- معايير وطنية National Norms
2- معايير محلية Local Norms
3- معايير متوسطات School Averages Norms
4- معايير الجماعات الخاصة Special Norm Groups

يُعد القسم الأول أكثرها شمولاً وانتشاراً، وأكثرها فائدة. أما النوع الرابع فإنه أقلها فائدة، وربما ينطبق هذا النوع على اختبارات المعلم. لذلك ينصح المعلم بحساب المتوسط الحسابي Mean والانحراف المعياري SD لدرجات كل مجموعة من الطلبة، وبعد أن يوزع الدرجات على الطلبة يعطيهم متوسط المجموعة وانحرافها المعياري، ويطلب منهم مقارنة درجاتهم مع هذا المتوسط، وبهذا الاجراء يكون قد حقق الحد الادنى من المواصفات المطلوبة في التقنين.

مراجع الفصل السادس

1- العاني، نزار (2003). القياس والتقويم وبناء الاختبارات المدرسية. الكويت: الجامعة العربية المفتوحة.

2- عبد الرحمن، سعد (2003). القياس النفسي - النظرية والتطبيق. القاهرة: دار الفكر العربي.

3- عبد الكريم، مجدي (1996)، التقويم والقياس في التربية وعلم النفس. القاهرة: مكتبة النهضة المصرية.

4- عدس، عبد الرحمن (1999). دليل المعلم في بناء الاختبارات التحصيلية. عمّان: دار الفكر للطباعة والنشر والتوزيع.

5- عودة، أحمد سليمان (2002). **القياس والتقويم في العملية التدريسية**. إربد: دار الأمل.

6- الكناني، ممدوح عبد المنعم وعيسى عبد الله جابر (1995). **القياس والتقويم النفسي والتربوي**. العين: مكتبة الفلاح للنشر والتوزيع.

7- نجيب، حسين علي؛ والرفاعي، غالب عوض (2006). **تحليل ونمذجة البيانات باستخدام الحاسوب: تطبيق شامل للحزمة SPSS**. عمان: الأهلية للنشر والتوزيع.

8- Allen, M.J., & Yen, W. M. (2002), **Introduction to measuement tehe-ory**. Prospects Heights 111: Waveland Press.

9- American Psychological Association (1999). **Standards for Educational and Psychological Tests**. Washington DC. APA.

10- Anastasi, A. & Urbina, S.(1982). **Psychological Testing (7thed.)**. Jersey: prentice Hall.

11- Chase, C. I. (1978). **Measurment for Educational Evaluation (2nd ed.)**. California: Addison-Weslru Publishing Company.

12- Cronbach, L. J. (1970). **Essentials of Psychological Testing**. New York: Harper and Row.

13- Ebel, R. 1. (1972). **Essentials of Educational Measurement**. NJ: Prentice - Hall.

14- George, D., & Mallery, P. (2003). **SPSS for Windows, Step by step; A simple Guide and Reference 11.0 Update (4thed)**. Boston: Allyn and Bacon.

15- Gronlund, N. E. (1976). **Constructing Achievement Tests**. New York: Prentice - Hall Inc.

16- Gronlund, N. & Linn, R. (1990). **Measurement and Evalnation in teaching (6th ed).** New York: Macmillan.

17- Gronlund, N. E (2006). **Assessment of Student Achievement (8th ed.).** Boston: Pearson Education, Inc.

18- Kublisyn, T. & Borich, G. (1987). **Edeucational Testing and Measurment, Classroom Application and Practice. (2nd ed).** Illinois: Scott, Foresman and Company.

19- Linn, R.L., & Gronlund, N. E (2000), **Measurement and Assessment in Teaching (7thed)**. New York; Macmillan Publishing Company.

20- McMahon, F. B. (1974). **Psychology: The Hybrid Science (2nd ed.).** NJ: Prentice - Hall.

21- Mehrens, W.A. & Lehman. J. L. (1991). **Measurement and Evaluation in Education and Psychology (4th ed.).** Florida: Harcourt Brace College Publishers.

22- Popham, W. J. (2000). **Modern Educational Measurement (3rd ed).** Boston: Pearson Education, Inc.

23- Stevens. S. (1951). **Mathmatic, Measurement, & Psychopgy**. In S. Stevense (Ed.). Handbook of Experimental Psychology. New York: Wiley.

24- Throndkle R. L., & Hagen, L. P. (1986). **Measurement and Evaluation in Psychology and Education.** New York: Macmillan Publishing Company.

25- Trice, A. (2000), **A Handbook of Classroom Assessment**. New York; Longman.

26- Wiersma, w. and Jurs, S. G. (1990) **Educational Measurment and Testing, (2nd ed.).** Boston: Ally and Bacon Inc.

الفصل السابع

اختبارات التحصيل وأنواعها

- العوامل المؤثرة في اختيار نوع الفقرات
- تحديد العدد المناسب لفقرات الاختبار
- اختبار المقال
- مزايا اختبار المقال
- الاختبارات الموضوعية
- إدارة وتصحيح الاختبار
- تحليل فاعلية فقرات الاختبار
- المعايير
- خطوات بناء الاختبارات التحصيلية

7

الفصل السابع

اختبارات التحصيل وأنواعها

بعد أن يُنجز المعلم أو مصمم الاختبار جدول المواصفات وفقاً للأسلوب الذي تطرقنا إليه في الفصل السابق تبدأ المرحلة الثانية المهمة في عملية بناء الاختبار، تلك هي صياغة فقراته طبقاً لما جاء في جدول المواصفات من حيث المحتوى والأهداف المطلوب قياسها. وقبل البدء بعملية صياغة الفقرات على المعلم أن يدرك إنه أمام أنواعٍ شتى من الفقرات التي قد تقيس نفس الهدف وتُشتَق من الموضوع الدراسي نفسهُ، وعليه أن يحدد النوع أو الأنواع التي سيلجأ اليها، وهل سيصمم كل الفقرات من نوع واحد أو ينوّعها؟ وهل هناك ظروف معينة تتعلق بطبيعة المادة الدراسية، أو بطبيعة المتعلم عليه أن يأخذها بالاعتبار قبل عملية الاختيار والتصميم؟

إن الجواب على هذا السؤال، نعم. ففي الجانب الأول على المعلم أن يدرك أن لكل نوعٍ من الفقرات ميزات ونقاط ضعفٍ وقوة وحالاتٍ يصلح في قياسها واخرى ليس بامكانه قياسها.

ويمكن توزيع تلك الأنواع في فئتين رئيسيتين هما:

1- النوع الأول: وفيه ينتج الطالب إجابة السؤال بنفسه وبإسلوبه الخاص ومنها: الاختبار المقالي، اختبار الاسئلة ذات الاجابات القصيرة, أسئلة الاكمال.

2- النوع الثاني: وفيه يختار الطالب إجابة السؤال من بين عدة إجابات معروضة عليه، ومنها: فقرات الصواب والخطأ, الاختيار من متعدد، فقرات المطابقة أو المزاوجة.

وسنقدم في الصفحات القادمة عَرضاً مفصلاً لكل نوعٍ من هذه الانواع، مُدعماً بأمثلة تطبيقية من مواد دراسية متنوعة.

أما بالنسبة للظروف التي يجب أن يراعيها المعلم عند اختياره لنوع فقرات الاختبار أو عددها او عوامل أُخرى مهمة، فيمكن إجمالها كما يأتي:

أولاً: العوامل المؤثرة في اختيار نوع الفقرات:

أ. **الهدف الذي نريد قياسه بالفقرة.** وهذا العامل ضروري ويجب على المعلم أخذهُ بنظر الاعتبار، إذ ليس جميع الفقرات تقيس كل الأهداف المراد قياسها، فهناك انواع يمكن استخدامها لقياس قدرة الطالب على التذكر أو معرفة الحقائق والأسماء والتواريخ وغير ذلك. وتكاد جميع أنواع الاختبارات التحصيلية الكتابية تقيس هذه القدرة سواء كانت التقليدية منها (المقالية) أو الموضوعية وبجميع أنواعها، لكنَّ هناك أهداف لا تصلح لقياسها الا أنواع محددة من الفقرات؛ فمثلاً اذا أراد المعلم قياس قدرة الطلبة على الإبداع والإبتكار من خلال الكتابة فإن الاختبارات المقالية تكاد تكون أفضل من غيرها في هذا المجال. وإذا أراد قياس قدرة الطالب على الفهم والاستيعاب فان اختبار الاختيار من متعدد، والصواب والخطأ الى حد ما قد يكونا أفضل الانواع الموضوعية لتحقيق هذا الغرض. وهكذا فإن على المعلم أن لا يتصور أن جميع انواع الفقرات تصلح لجميع الأهداف المراد التأكد من تحققها في سلوك الطالب، ولذلكَ وجبَ على المعلم أن يكون مُلماً بمعرفة طبيعة كل نوع من الاختبارات وأي الأهداف يصلح لقياسها، وإذا لم يكن كذلك فان عليه مُراجعة الكتب والمصادر التي تساعده وهي كثيرة وفي اللغتين العربية والانجليزية.

ب- **طبيعة محتوى المادة الدراسية:** أما العامل الثاني الذي يجب على المعلم أو مصم الأختبار الاهتمام به فهو محتوى المادة الدراسية، أو الموضوعات المطلوب تغطيتها بالاختبار، وهي بلا شك تختلف في طبيعة محتواها وأسلوب عرضها في الكتاب المدرسي، ولأن لكل مادة دراسية طريقة تدريس تناسبها وقد لا تناسب جميع المواد الأخرى، فإن بعض هذه المواد أو الموضوعات تصلح لقياس أهداف سلوكية معينة وقد لا تصلح لقياس

أهداف اخرى. فموضوعات مادة الرياضيات مثلاً تصلح جداً لتصميم فقراتٍ لقياس قدرة الطالب على مهارة التطبيق بشكل أكبر وافضل من قياسها لمهارتي التذكر أو التركيب. أما مادة التاريخ فهي مجالاً خصباً يمكن للمعلم أن يصمم منها فقرات لقياس القدرة على التحليل أو التذكر، وهكذا بالنسبة للمواد الدراسية الاخرى. والمعلم المدرك لهذا العامل سيكون بلا شك أفضل من غيره في بناء إختبارٍ تحصيلي جيد.

ج- خبرة المعلم ومهارته في تصميم إختبارات التحصيل: مما لا شك فيه إن المعلمين يتباينون في مهارة صياغة الاختبارات التحصيلية، فمنهم من سبق وأن تخرج من كليات التربية ودرس موضوعات القياس والتقويم وكيفية إعداد الاختبارات وهؤلاء قد يكونوا أكثر قدرةً على أداء هذه المهارة، وقد نَجد على العكس من هؤلاء من تنقصهم الخبرة في مجال صياغة الاختبارات لا سيما وان بعضهم لم يتخرج من كلّيات التربية، أو لم يسبق له أن درس مقررات متخصصةً في القياس والتقويم مما ينعكس سلباً على قدرته في بناء إختبارٍ تحصيلي جيد، الا ان ذلك يجب ان لا يمنع مثل هؤلاء من متابعة ما تتضمنه الكتب المؤلفة في مجال تطوير الاختبارات لأكتساب هذه المهارة أو الالتحاق بدورات وورش العمل التي تقيمها وزارة التربية في مجال تصميم الاختبارات أو الاطلاع على بعض الاختبارات المنشورة في المجالات العلمية أو على مواقع الانترنت في الوقت الحاضر.

ثانياً: تحديد العدد المناسب لفقرات الاختبار

إن تقدير العدد المناسب لفقرات الاختبار أو تحديد طول الاختبار يخضع لعوامل عديدة منها:

أ- **نوع الفقرات المستخدمة**: أي هل فقرات من النوع المقالي الذي يعتمد على توفير الاجابة من قبل الطالب والتي تحتاج الى جهد كتابي وتعبيري ووقت أكثر من غيرها من الفقرات الاخرى؟ وهنا على المعلم أن يكون

مُقّـلاً في عدد الفقرات . أما اذا إختار فقرات من النوع الموضوعي الذي يعتمد على الاختيار فإن بامكانه زيادة عدد الفقرات على أن لا تتجاوز ما مسموح بهِ من الوقت المخصص للاختبار . وقد بينت الدراسات الميدانية ان الطالب المتوسط المستوى يحتاج الى (30-40 ثانية) لقراءة واستيعاب فقرة الاختيار من متعدد ومن (5-15) دقيقة لاجابة فقرة من نوع المقال متوسط الصعوبة .

ب- **عمر الطالب ومستواه التعليمي:** هناك فقرات إختبارية تتناسب مع الأطفال ومنها فقرات المطابقة أو المزاوجة لسهولة إجراء المطابقة من قبل هؤلاء الأطفال ، او اختبار الاختيار من متعدد بثلاثة بدائل وحتى اختبار الصواب والخطأ يمكن استخدامها مع تلاميذ المرحلة الابتدائية ، يضاف الى ذلك الاسئلة المقالية محددة الاجابة لا سيما المتعلقة بتذكر الأسماء والتواريخ والمواقع . اما في المراحل المتقدمة كالثانوية والجامعة فيمكن إستخدام الاختبارات المقالية التي تقيس التحليل والنقد والابتكار والاختيار من متعدد التي تقيس الفهم والتحليل والتركيب والتقويم لان الطالب سيكون بمستوى عمري وتعليمي يؤهله لإدراك المطلوب منه في هذا النوع من الاختبارات .

ج- **طول الفقرة ودرجة تعقيدها:** إذا كانت الفرات تعتمد اجاباتها على اشكال بيانية او احصاءات معينة ، أو تعتمد على نصٍ مكتوب طويل وجب على مصمم الاختبار أن يقلل من عدد الاسئلة بِشكلٍ أقل من فقرات تعتمد على فراغات أو إختيارٍ من عددٍ من البدائل .

د- **نوع العمليات العقلية المراد قياسها:** فالفقرات التي تقيس الجوانب المعرفية التي تعتمد على التفكير العميق كالنقد والتحليل فيمكن إقلال عددها في الاختبار لأنها تتطلب زمنا أطول للإجابة عليها .

هـ- **مقدار ما تتطلبهُ الفقرة من عمليات حسابية أو معالجات كمية.** تتطلب الفقرات ذات العمليات الحسابية الكمية جهداً ووقتاً أطول من الطالب

سواء كانت باسلوب مقالي أو حتى موضوعي كما هو الحال في إختبارات الرياضيات، لذلك على المعلم أخذ هذا الأمر بنظر الاعتبار في تصميم الاختبار بحيث يعمل موازنة بين الوقت وطول الاختبار.

وبعد هذه المقدمة التي توضح ضرورة الاهتمام ببعض العوامل المهمة عند البدء بصياغة فقرات الاختبار التحصيلي، لا بد من إطلاع المعلم أو مصمم الاختبار على الأنواع الخاصة بقياس التحصيل المعرفي للطالب من حيث طبيعة كل نوعٍ واستخداماته، ونقاط ضعفه، وجوانب القوة فيه، والقواعد العلمية المطلوب مراعاتها في صياغة تلك الانواع، سواء كانت التقليدية منها (المقالية)، أو الحديثة (الموضوعية) وهذا ما سيتم التطرق له بالتفصيل في الصفحات القادمة.

إختبار المقال Essay Examination

سادَ هذا النوع من الاختبارات الكتابية في ميدان تقويم تحصيل الطلبة لفترة طويلة من الزمن ولا يزال يستخدم في العديد من المؤسسات التعليمية وبكافة مراحلها لعدة أسباب لعل أهمها سهولة اعداده من قبل المعلمين والتاريخ الطويل لاستخدامه (Mehrens & Lehman, 1991) وتشير بعض الادبيات الى أن هذا النوع وجد في التربية الصينية القديمة قبل الميلاد. أما في العصر الحديث فيشير (Chase,1978) الى أن هذا النوع سيطر على عملية القياس الصفي خلال الثلث الأول من القرن العشرين سيطرة تامة قبل أن توجه له حملة نقد قوية ادت إالى إعطاء المجال لظهور حركة الاختبارات التحصيلية الموضوعيّة التي قامت على إبراز نقاط ضعف ومساويء الاعتماد الكلي على اختبارات المقال كأسلوب لقياس وتقويم تحصيل الطلبة.

وقد سُميت الاختبارات المقالية بهذا الإسم لأنها في أغلب الأحيان تتطلب من الطالب كتابة إجابته بما يشبه كتابة المقال. ويشمل هذا النوع مدىً واسعاً من الاجابة يتراوح من كَتابة جملة بسيطة مثل: اكتب جملة تستخدم فيها الفعل (قرأ)، الى اسئلة تتطلب إجابة طويلة تتطلب كتابة صفحات عدة.

وتبدأ أسئلة هذا النوع عادة بكلمات مألوفة للطالب مثل: إشرح، ناقش، علل، تكلم عن، عرف ما يأتي، قارنً، بين أسباب . . . إلخ.

وعلى الرغم من أن الاختبارات المقالية هي من نوع الاختبارات التي تسمى بـ (تجهيز الإجابة Supply Type) سواء كانت الاجابة قصيرةً أم طويلة، إلا إن هناك من يصنفها الى نوعين (Gronland,1985, Popham, 1981) هما:

أ- المقالي ذو الاجابة المحددة Restricated response وفيه يطلب من الطالب إجابة محددة وعدم الخروج عن مجال السؤال. ويصلح في قياس الجانب المعرفي لا سيما في المراحل الدراسية الدنيا. مثال: أذكر الفروق بين الاعراب والقواعد في اللغة؟

أو: حدد أهم أربعة أسباب أدت الى انهيار (ألمانيا في الحرب العالمية الثانية)؟

ب- المقالي ذو الاجابة المفتوحة Extended response وفيه لا يحدد الطالب بقيود معينة بل تترك له حرية الكتابة بما يعتقد أنه يمثل الاجابة الصحيحة. وكثيراً ما يستخدم هذا النوع في المراحل الدراسية العليا لقياس نواتج متقدمة من التعلم تمثل القدرات العليا لدى الطالب.

مثال:

من خلال دراستك لكل من التقويم الكمي والتقويم النوعي بين أوجه الشبه والاختلاف بين النوعين وأيهما يصلح أكثر للإستخدام في تقويم البرامج التربوية، أسند رأيك بحجج علمية وموضوعية؟

وخلاصة القول: ان الاختبار المقالي هو أي اختبار يطلب فيه من الطالب صياغة افكاره باجابة مكتوبة سواء كانت تلك الأجابة طويلة أو قصيرة، محددة أو غير محددة.

مزايا إختبار المقال:

1- اذا تم تصميم فقرات هذا النوع من الاختبارات فإنها يمكن أن تقيس مستويات عالية من القدرات كالأبداع وربط وتحليل ومقارنة المعلومات

ببعضها ووضع استنتاجات تدل على امكانية الطالب على تقديم الاجابة المطلوبة بشكلٍ جيد وقد حدد بعض الباحثين في مجال التقويم والقياس ومنهم (Kubiszyn & Borich, 1987; Wiersma & Jurs, 1990) وعدداً من المهارات أو القدرات التحصيلية الممكن تقويمها في اختبارات المقال ومنها:

تحليل العلاقات بين الافكار أو المفاهيم العلمية، ترتيب العناصر وفق نسق معين، اجراء المقارنات بين أفكار او عمليات معينة، وضع واطلاق افتراضاتٍ معينة، وضع خلاصات واستنتاجات معينة، تحديد العلاقات بين الأسباب والنتائج، صياغة فرضيات، تنظيم بيانات ومعطيات خاصة لدعم وجهة نظر معنية، للتوصل الى حلول تتعلق بمشكلة معينة، تقويم جودة او قيمة فكرة أو منتج أو عمل معين، خلق حلٍ أصيل مبدعٍ لمشكلة محددة، توضيح كيفية تطبيق لمبدأ سيكولوجي معين، وغير ذلك.

2- في هذا الاختبار يأخد الطالب حريةً أكثر عند الاجابة للتعبير عن نفسهُ وبذلك تكون الفرصة سانحة أما المعلم لتحديد معالم شخصية الطالب وقدراته العلمية ومقارنة مستويات الطلبة ببعضهم بما ينعكس إيجابيا على موضوعية تقويمهم وفقاً لما يقَدمونهُ في تلك الإِجابات. يتميز هذا الاختبار بتنوع محتوى إجابات الطلبة على فقراته.

ان تنوع الاجابات في هذا الاختبار تعود الى عدة أسبابٍ منها:

أ- إن السؤال قد يكون واسعاً شاملاً وغير محدد، مما يدفعً الطالب الى كتابة ما يتصور انه يمثل الاجابة الصحيحة.

ب- قد يكون السؤال غامضاً ويحمل تفسيرات مختلفة وبالتالي فإن كل طالب يحدد الاجابة وفقاً لتفسيره الخاص للسؤال. فعلى سبيل المثال لو سأل المعلم طلابه "ما هي أهم أسباب الحرب العالمية الثانية؟" فإن اجابات الطلبة بلا شك ستتباين، وسيبدأ كل منهم الاجابة من النقطة التي يعتقد انها تمثل أهم الأسباب التي سبقت الحرب وأدت إلى نشوبها. فقد يعزي أحد الطلبة نشوب الحرب إلى مقتل ولي عهد النمسا، وقد يُعزيها

آخر للهجوم على باخرة تعود لدولة معينة. وقد يعزوها آخر الى نشوب نزاعٍ مسلح على حدود دولتين متجاورتين. ولعل هذه التباينات في الاجابات تعود الى غموض السؤال وعدم تحديده بنقاطٍ معينة ينطلق منها الطلاب للاجابة فتقل فجوة التباعد في اجاباتهم.

3- في هذا الاختبار تقلّ أمام الطلبة فرصة التخمين Gussing التي كثيراً ما تلازم الاختبارات الموضوعية، فالطالب في هذا النوع يكتب اجابات مطولة يصعب تخمينها لانها لا تمثل وضع كلمات في فراغات او إنتقاء بديل من مجموعة بدائل أو وضع اشارات معينة كما يحدث في اجابة الاختبارات الموضوعية.

4- إن تصميم فقرات هذا النوع أسهل على المعلم مقارنةً بالاختبارات الموضوعية، ولذلك يميل لها المعلمون إقتصاداً في الجهدِ والوقت.

5- تعتمد اجابة الطالب في هذا النوع على قدرتهِ على التعبير اللغوي. فنفس المعلومات والافكار يمكن أن تعرض بإسلوب جَذاب وشيق من قبل طالب معين بينما تُعرض بإسلوب جاف وركيك وغير شيق من قبل طالب آخر.

6- ان استخدام المعلم لهذا النوع من الاختبارات بشكل متكرر يحدد طريقة مذاكرة الطلبة، حيث يدفعهم الى التركيز على فهم المادة واستيعابها من جهة، وحفظها من جهة اخرى، يضاف الى ذلك ان هذا النوع سيدفع الطالب أيضاً الى الاهتمامٍ بقدرته على التعبير المنظم إثناء الكتابة ووضع افكاره بشكل متسلسل ودراسة المادة بشكل معمق وليس سطحياً مما يؤدي الى تطور هذه المهارات الجيدة لدى الطالب.

نقاط الضعف في اختبار المقال:

على الرغم من المزايا الجيدة التي يتمتع بها الاختبار المقالي فان هناك بعض العيوب أو نقاط الضعف التي يتصف بها هذا الاختبار ومنها:

1- لا يمكن أن يحقق هذا الاختبار مستوىً جيداً من الصدق بسبب قلة عدد فقراتهِ وما ينتج عن ذلك من عدم تغطيتها أو تمثيلها لمحتوى المادة

الدراسية من جهة وأهداف تعليمية كافية من جهة اخرى، إذ كثيراً ما يحصر المعلم الاسئلة في عدد محدد من وحدات يعتقد انها مهمة ويترك الوحدات الاخرى وعدد محدود جداً من الأهداف وبذلك لا يكون الاختبار صادقاً في تمثيله لمحتوى المادة الدراسية والأهداف التعليمية بالشكل المطلوب.

2 - ضعف موضوعية وثبات تصحيح إجابات الطلبة على هذا الاختبار من قبل المعلم بسبب تأثر عملية التصحيح بالعوامل الذاتية للمصحح والظروف المحيطة به وعوامل اخرى كتأثير الهالة Halo Effect أي الانطباع الذي سبق للمعلم أن كونه عن الطالب سالباً أم موجباً. ولا يقتصر ضعف الموضوعية على اختلاف مجموعة المصححين في تحديد الدرجة التي تستحقها إجابة الطالب كما وجد من بحوث عديدة أجريت في هذا المجال (ديمرجي والآلوسي، 1980)، بل إن المصحح نفسه قد يتناقض مع نفسه. فقي احدى الدراسات قدّم أحد الباحثين لمعلم واحد عشرة أوراق إجابة عن اسئلة مقالية لعشرة طلاب وقام المعلم بتصحيحها. وبعد شهرين أعيدت نفس الاوراق (مستنسخة) ثانيةً الى نفس المعلم الذي قام بتصحيحها فوُجدَ أن ذلك المعلم تناقض مع ما وضعه من درجات على ذات أوراق الإجابة فمرةً يُنقص العلامة التي وضعها في المرة الاولى، وأُخرى يزيدها، وقد وجد أن الفروق بين تصحيح بعض الاوراق في المرة الاولى والمرة الثانية بلغت (25%) أي الى حوالي ربع الدرجة الكلية، وان معامل الارتباط بين علامات التصحيح الاول والتصحيح الثاني بلغ (0.45) ويمثل علاقة ضعيفة (Tiegs,1952).

3 - تتأثر درجة الطالب في هذا الاختبار ببعض الأمور التي لا علاقة لها بالاهداف التي يقيسها مثل جودة خط الطالب وتنظيمه لورقة الاجابة وقدرته الانشائية عند الكتابة مما يضعف موضوعية التقويم الذي يفترض أن يُركز على قدرة الطالب على الاجابة والمخرجات التي يهدف الاختبار الى قياسها.

4 - تحتاج عملية تصحيح اجابات الطلبة الوقت والجهد الكبير لما تتطلبه هذه العملية منه قراءة متأنية لأوراق الإجابة، وتزيد هذه الحالة في إرهاق المعلم ومللهُ عندما يكون عدد الطلاب كبيراً وتتطلب الاسئلة إجابات مطوّلة.

5 - تدخل عامل الصدفة أو الحظ في حصول الطالب على درجة عالية أو منخفضة وذلك بسبب صياغة الاسئلة من وحدات أو موضوعات معينة من المادة الدراسية، فالطالب الذي ركز بالصدفة على هذه الموضوعات قد يحصل على درجة عالية، وبالعكس منه الطالب الذي ركز اهتمامه على موضوعات اخرى بدلاً من تلك الموضوعات قد يحصل على علاقة متدنية في ذلك الامتحان. كما أن ذلك قد يُضعف موضوعية مقارنة مستوى الطالبين والحكم عليهما بشكل دقيق وفقاً لنتيجتيهما في ذلك الإمتحان.

بعض القواعد والارشادات المهمة في استخدام اختبارات المقال

بناءً على ما تمت الاشارة له من الجوانب الإيجابية ونقاط الضعف في الاختبارات المقالية، ومن أجل تقليص مساحة الجوانب السلبية في هذا الاختبار والاستفادة منه في تقويم الأهداف المهمة عالية المستوى لا بد من تقديم بعض الارشادات والقواعد العلمية التي يستفيد منها المعلمين والباحثين. وستقدم هذه الارشادات والقواعد في فئتين تتعلق الأولى بتصميم فقرات المقال، بينما تتعلق الثانية بقواعد تصحيح إجابات الطلبة على أسئلة المقال.

أولا: القواعد المتعلقة بتصميم فقرات المقال:

1 - بشكل عام يؤكد العديد من الباحثين على ضرورة حصر استخدام هذا النوع من الاختبارات على المخرجات التعليمية التي لا يمكن تغطيتها بشكل كافي في الانواع الاخرى من الاختبارات كالتحليل، والاحكام والتفكير الناقد والربط والاستنتاج وهذه الأهداف يجب أن تكون واضحة في ذهن المعلم قبل أن يفكر باختيار اختبار المقال. أما تقويم الجوانب المعرفية التي لا تتطلب اكثر من تذكرها وإسترجاعها كالاسماء والتواريخ

وغيرها فيمكن تقويمها بالاختبارات الموضوعية (Wiersma & Jurs,1990) وبذلك نختصر الجهد الذي يبذله الطالب أثناء الكتابة في الاختبار المقالي وكذلك الوقت.

2 - من الضروري جداً أن يكون السؤال المقالي واضحاً ومحدداً يدرك فيه الطالب ما يجب عليه القيام به والابتعاد كلما أمكن عن الغموض والإطناب الذي يمكن أن يجعل الكثير من الطلبة يتخبطون في إجاباتهم أو يكتبوا كل ما يقع في ذاكرتهم عن الموضوع الذي أُشتقَ منه السؤال وفي ذلك هدراً للجهد والوقت وإعطاء الفرصة للعوامل الذاتية لتلعب دورها بما يؤدي الى اختلال عملية التقويم العادل والدقيق للطالب. فلو تضمَّنَ اختبار المقال سؤالاًمن مثل "فيما يخص الحد من إنتشار مرض السل الرئوي ناقش المؤسسات التي تساهم في الصحة العامة للمجتمع" نلاحظ أن هذا السؤال عاماً غامضاً لا يمكن أن يتفق إثنان من الطلبة على إجابته، فما المقصود بكلمة (ناقش)؟ هل المقصود بها نقد أو تقويم، أم مسائلةَ، أم تحديد النقاط الايجابة والسلبية في عمل تلك المؤسسات؟ أم ذكر ما تقدمهُ تلك المؤسسات لرفع المستوى الصحي للمجتمع؟ كما يمكن أن نتساءل ايضاً عما يقصده واضع السؤال بكلمة (تساهم في الصحة العامة للمجتمع). هل يريد من الطالب ان يتطرق لذكر ما تقدمه تلك المؤسسات من توعية صحية فقط؟ أم ما تقدمه من العلاج المباشر للمرضى؟ أم مساهمتها في توجيه الأفراد للوقاية من الأمراض، أم ما تساهم به في دعم البحوث العلمية الخاصة برفع المستوى الصحي للمجتمع؟ هذه اسئلة كثيرة قد يُثيرها الطالب المتفوق مع نفسه ويُثيرها بعضاً منها الطالب المتوسط المستوى، وقد يلجأ بعض الطلبة الى عملية التخمين. ويبقى تساؤلاً آخر يوجه لواضع السؤال ألا وهو: من أين تريد الطالب أن يبدأ إجاباته؟ وعلى ماذا يُركز؟ ثم كيف ستحدد العلامة التي يستحقها كل طالب، وما هو معيارك لتحديد العلامات التي تعكس الفروق في مستويات الطلبة؟

وهنا لا بد أن نوجه واضع السؤال الى تجنب مثل هذه الصياغة عند وضعه للأسئلة، وأن يقرأ السؤال عدة مرات، ويسأل نفسه عما يريده من الطالب أن يظهر من قدرة أو سلوك أولاً، وما هي حدود السؤال؟، وما هي متطلبات الاجابة وكيف سيوزع الدرجة على عناصر الاجابة وفقاً لأهميتها؟ وانطلاقاً من هذه التساؤلات لا بد من إعادة صياغة هذا السؤال ليكون على الشكل اللآتي:

"ما هي برأيك الخدمات التي يمكن أن تقدمها كلٍ من المؤسسات الصحية التالية للتخفيف من انتشار مرض السل الرئوي في الجانبين الإعلامي والعلاجي؟".

أ- المؤسسات الصحية المركزية.

ب- المؤسسات الصحية المحلية.

ج- المؤسسات التعليمية.

وهنا نُلاحظ أن السؤال أصبح محدداً وذا عناصر واضحة، وان المهمة التي على الطالب أن يقوم بها أصبحت واضحة وتتحدد في جانبين من عمل بتلك المؤسسات لتخفيف إنتشار المرض، وهما: الجانب الاعلامي والجانب العلاجي. وبالاضافة الى ذلك فان عملية التصحيح ستكون أكثر دقةً وموضوعيةً. فلو فرضنا ان ما يستحقه هذا السؤال هو (15) درجة؛ فبامكان المعلم أن يوزع تلك الدرجات بعدالة على العناصر الثلاثة، أي تخصيص (5) درجات لكل عنصر، وبذلك نقلل من العوامل الذاتيةونزيد من العوامل الموضوعية في تحديد الدرجة، وبإمكان المعلم أن يحدد أية طريقة يجدها مناسبة، بعد أن يحدد عناصر الاجابة ويفرعها حسب خبرته.

3 - إبدأ اسئلة المقال بكلمات أو عبارات من مثل: قارن، أعط السبب في.. بين الفروق، أعط أمثلة حقيقية لـ..، تنبأ بما يمكن أن يحدث عندما، أنقد، حيث أن مثل هذه الكلمات والعبارات تقترن في العادة بأفكار جديدة ومهمات تتطلب من الطالب أن يختار أفكار رصينة من معرفته المخزونة

وينظمها ليوظفها في الاجابة. وقلل من بدأ السؤال المقالي بكلمات شائعة من مثل: ماذا، مَن، متى، عدّد، عرّف. لأن هذه الكلمات تتطلب عادةً إستدعاء معلومات ومهمات تعتمد على الحفظ والاستظهار ويمكن اللجوء إلى ذلك عندما نريد قياس هذه المهارة لدى الطالب.

مثال ضعيف: عدد ثلاثة فقط من أعضاء الجهاز الهضمي عند الانسان؟

مثال جيد: بعد عشر سنوات من الحرب في فيتنام سحبت أمريكا قواتها عام 1965، تنبأ بما كان يمكن أن يحدث لو أنها لم تسحب تلك القوات في التاريخ أعلاه، وأبقت على نفس مستوى قواتها الذي دخلت به الحرب؟

4- تجنب اعطاء الطلبة حرية أختيار بعض الأسئلة وترك أسئلةً أخرى، كأن تطلب منهم إجابة أربعة أسئلة من مجموع ستة، لان ذلك يقلل من صدق الاختبار من جهة، ويقلل من فرص وموضوعية المقارنة بين اداء الطلبة. يضاف الى ذلك ان بعض الطلبة عندما يعرفون ان هناك فرصة للاختيار سوف لن يركزوا على جميع أجزاء المادة الدراسية.

5- حاول أن توازن بين عدد أسئلة الاختبار والوقت المخصص له من أجل أن تساعد الطلبة على الاجابة دون التفكير أو القلق الناشئ عن ضيق الوقت وما يمكن أن ينشأ عن ذلك من السرعة دون الدقة في الاجابة، ويمكن معالجة الموضوع بتحديد طول الاجابة الخاصة بكل سؤال بإستخدام عبارات أو توجيهات من مثل: على أن لا تزيد اجابة السؤال عن صفحتين، او: انقد بما لا يزيد عن صفحة واحدة موضوع...، أو اترك مساحة محددة للأجابة عن كل سؤال على نفس ورقة الأسئلة.

6- حاول أن تضمّن الاختبار المقالي عدداً من الاسئلة (الفقرات) التي تتطلب إجابات قصيرة ومحددة لأن ذلك يمكن أن يُزيد من شمولية الاختبار وتمثيله للمحتوى الدراسي، ويقلل من تأثير العوامل الذاتية للمعلم كما في الاسئلة التي تتطلب إجابات مطوّلة، إضافةً الى أن هذا الاسلوب يساعد المعلم على قراءة الاجابات المحددة بجهد ووقت أقل وتركيز أكبر، وأخيراً إن هذا الأسلوب يوجه الطالب نحو الإجابة المطلوبة بشكل أدق وأسرع.

ثانياً: القواعد المتعلقة بتصحيح اجابات أسئلة المقال

أن العوامل التي تؤثر على صدق وموضوعية نتائج إختبار المقال لا تنحصر فقط في أساليب صياغتها، بل تمتد الى عملية تصحيح إجابات هذا الاختبار ولعل أبرزها العوامل الذاتية للمصحح مما دفَع الباحثين في مجال القياس الى تقديم العديد من القواعد والإرشادات الهامة التي على المعلم أخذها بنظر الاعتبار حتى تؤدي الى زيادة موضوعية التصحيح وتجعل الاختبار يتمتع بدرجة أعلى من الصدق والثبات .(Gotman, 1978) ومن تلك القواعد ما يأتي:

1- من الضروري جداً وضع مُخطط لعملية التصحيح من قبل المصحح قبل البدء بها، ويمثل هذا المخطط العناصر الأساسية للإجابة التي تتطابق مع عناصر السؤال، وتوزيع الدرجات على العناصر وفقاً لأهميتها، ويعتبر محكاً يطابق به المعلم إجابات طلابه. ويفضل أن يقوم المعلم قبل وضع هذا المخطط بالاطلاع على عينة من إجابات الطلبة على ذلك السؤال ليكون على علم بالاتجاه العام لتلك الاجابات ويأخذها بنظر الاعتبار عند توزيع الدرجات. لنفرض أن احد فقرات المقال ينص على ما ياتي: "ما المقصود بمفهوم " التعلم الشرطي" الذي أطلقهُ بافلوف؟ وما هي إجراءات التجربة التي أجراها للتوصل الى هذا المفهوم؟ إستنتج ثلاثة تطبيقات تربوية يمكن للمعلم إستثمارها في عمله التربوي وفقاً لمنطلقات هذا المفهوم؟" .

وهنا نلاحظ أن على الطالب أن يركز إجابتهِ على هذا السؤال في ثلاثة عناصر هي:

أ- معنى مفهوم التعلم الشرطي.

ب- إجراءات التجربة الخاصة بالتعلم الشرطي.

ج- إستنتاج ثلاثة تطبيقات تربوية.

ولنفرض أن المعلم خصص لهذا السؤال (20) درجة، فبامكانه على سبيل المثال أن يوزعها كالآتي:

أ- (5) علامات.

ب-(6) علامات.

ج- (9) علامات. موزعة على التطبيقات الثلاثة بالتساوي. ولا بد لهذا الأجراء أن يُقلل من أثر العوامل الذاتية، ويّقرب عملية التصحيح والتقويم من الموضوعية وإن لم تكن الموضوعية التامَّة.

2- من المفضل للمعلم أثناء عملية تصحيح اجابات الطلبة عدم اطّلاعه على أسمائهم تحاشياً لتأثير الهالة أو العوامل الذاتية الإخرى، ولا بأس لَلمعلم أن يطلب من طلابه كتابة أرقام تسلسلاتهم الصفية بدلاً من أسمائهم.

3- من المفضل تصحيح إجابة كل سؤال على حدة ولجميع أوراق الطلبة، بدلا من تصحيح جميع الاسئلة في الورقة الواحدة الخاصة بكل طالب، فالأسلوب الأول له عدة فوائد منها:

أ- تقليص اثر الهالة التي قد تنتقل من سؤال الى آخر كما في الاسلوب الثاني. فمثلاً لو صحح المعلم اجابةً الطالب على السؤال الاول ووجدها ممتازة فانه سوف ينقل تأثره الايجابي عن تلك الاجابة الى إجابة السؤال الثاني الذي لم يكن بالمستوى المطلوب، وقد يتغاضى عن اخطاء الطالب في هذا السؤال ويعطيه درجة لا يستحقها، وقد ينطبق نفس الشيء عندما تكون اجابة الطالب على سؤال سابق ضعيفة جداً فينتقل تأثره السلبي الى إجابة ذلك الطالب علىً السؤال التالي، فيعطيه درجة قليلة رغم إجابته الجيدة.

ب- يساعد هذا الاسلوب على اتباع نسق واحد في عملية التصحيح لجميع أوراق الطلبة.

ج- يساعد المصحح على تركيز أنتباهه في موضوعٍ واحد، وهو المتمثل بالاجابة وعدم تشتته بين موضوع وأخر من الموضوعات المختلفة التي تتضمنها اجابات الاَسئلة.

د- ان هذا الاسلوب يساعد المعلم على اجراء مقارنات بين مستويات طلابه في موضوع واحد، وتحديد الصعوبات التي بتلمسها في

اجاتاتهم عن ذلك السؤال مما يساعد في اعادة النظر بطريقة تدريسه لذلك الموضوع بعد مناقشة الموضوع معهم في الصف وبذلك يطبق مبدأ (التغذية الراجعة).

4- يفضل قيام المصحح بتغيير ترتيب أوراق اجابات الطلبة بعد تصحيح كل سؤال لجميع الاوراق؛ حتى لا تتأثر الدرجة التي يستحقها طالب معين بإجابة الطالب الذي يسبقهُ سلباً أو ايجاباً في كل مرة. فلو فرضنا أن ترتيب أوراق إجابات الطلبة كان كالآتي: (محمد، أحمد، محمود، مصطفى، حازم، عبد الله)، وكان الطالب (محمود) ضعيف في جميع إجاباته على أسئلة الامتحان، فمن المحتمل أن المصحح يضجر كلما يصحح اجابة سؤال لهذا الطالب، وربما تنعكس حالة ذلك المصحح في كل مرة على إجابة الطالب الذي يليه وهو (مصطفى) رغم انها إجابة جيدة، وبدلاً من أن يشمل هذا الاثرَ طالباً بعينه، فان اعادة ترتيب الاوراق عقب تصحيح كل سؤال سيوزع هذا الاثر على أكثر من طالب وليس على طالب معين.

5- إذا كان عدد الاوراق المطلوب تصحيحها كبيراً؛ فيفضل أن يأخذ المصحح فترات راحة قصيرة بعد الانتهاء من تصحيح مجموعة محددة من الاوراق منعاً للتعب والإجهاد المتواصل الذي يمكن أن يؤثر سلباً على دقة عملية التصحيح لا سيما في تصحيح الاوراق التي يكون تسلسلها في أواخر المجموعة.

6- إذا كان الاختبار المقالي يتوقف على نتائجه اتخاذ قرار مهم يخص الطلبة؛ فيفضل الاستعانة بأكثر من مصحح، ويؤخذ متوسط تقديراتهم لكل طالب توخياً للدقة وتقليص دور العوامل الذاتية للمصحح الواحد.

الإختبارات الموضوعية Objective Tests

ظهرت هذه الاختبارات نتيجةً للانتقاد الموجه للاختبارات المقالية بسبب نقاط ضعفها في جوانب عديدة سبق أن أوضحناها في الصفحات السابقة. لقد سميت هذه الاختبارات بـ (الموضوعية) لعدم تأثرها بالعوامل الذاتية إذا ما

أحسن تصميمها بشكل جيد. وتعني الموضوعية أيضاً (الاتفاق في الاحكام) اي ان عدد المصححين مهما اختلف فانهم يتوصلون إلى نفس الحكم وتحديد نفس الدرجة التي يستحقها الطالب في إجابتهِ على فقرات هذه الاختبارات.

لا يتطلب هذا النوع من الفقرات من الطالب كتابة جملٍ طويلة كما في إختبارات المقال، بل يُفترض بهِ وضع إشارات معينة أو أرقاماً او كلمات أو حروف في أماكن محددة. وتُستَخدم هذه الاختبارات بشكل واسع الانتشار في الدول الغربية وعلى الخصوص الولايات المتحدة الامريكية حيث تتوفر فيها مؤسسات خاصة تقوم ببناء وطبع ونشر هذه الاختبارات نتيجةً للمعرفة العلمية المتراكمة في بنائها واستخدامها وتوفر أجهزة الطباعة والتقنيات الحديثة التي سرَّعت في انتشارها، بالاضافة إلى خبرة المعلمين في كيفية بنائها وإدارتها وتصحيحها لا سيما في بعض انواعها العديد من الاهداف التعليمية المعرفية وغير المعرفية كما سنأتي على ذلك لاحقاً.

ميزات الفقرات الموضوعية:

إن أهم ما تمتاز به هذه الفقرات هي أنها:

أ- لا تتأثر بالاحكام الذاتية للمصحح.

ب- تتمتع بدرجة جيدة من الصدق والثبات في أغلب أنواعها.

ج- تغطي أغلب أجزاء المادة الدراسية نظراً لكثرة عدد الفروع فيها.

د- لا تحتاج الى جهدٍ كبير في تصحيح إجاباتها لاسيما بعد إنتشار أجهزة الحاسوب.

هـ- غير مرهقة للطالب من حيث الحفظ الأصم أو الكتابة.

و- تقيس مستويات عالية من القدرات والأداء في بعض أنواعها مثل: الفهم والاستيعاب، التطبيق، التحليل والتفسير.

ز- تستخدم مع جميع مستويات الطلبة من المراحل الدنيا وحتى المراحل الدراسية العليا.

نقاط ضعف الفقرات الموضوعية:

على الرغم من الجوانب الايجابية التي تتميز بها هذه الففرات لكنها - شأن كل ادوات القياس في العلوم السلوكية- تتصف بنقاط ضعف كثيراً ما يردد بعضها أولئك الذين ينتقدونها، ومنها:

أ- تتطلب من المعلم جهداً ووقتاً كبيرين وخبرةً وافية لتصميمها.

ب- امكانية إستخدام الطالب لعملية (التخمين، Guessing) واحتمال الحصول على درجات عالية رغم عدم إستعداده للإختبار.

جـ- يصعب استخدامها لتقويم بعض المستويات المتقدمة من القدرات كالابداع والتحليل، والتعبير عن الذات من قبل الطالب.

د- سهولة استخدام (الغش Cheeting) من قبل الطالب للحصول على الاجابة الصحيحة من خلال النظرة السريعة الى ورقة اجابة طالب آخر يجلس بجواره.

هـ- يتطلب اعداد هذه الاختبارات إمكانات طباعية وفنية مكلفة مادياً وتستهلك كميات كبيرة من الورق.

أنواع الفقرات الموضوعية. Types of objective items

يمكن تصنيف جميع الفقرات الموضوعية في ثلاث فئات هي:

أ- تجهيز الإجابة Supply type ومنها فقرات "الاجابات القصيرة (Short an-swer) أو الاكمال، أو الفراغات كما يطلق عليها من بعض الطلبة.

ب- إختيار الإجابة Select type ومنها: فقرات الإختيار من متعدد. Multi-ple- choice فقرات المطابقة (المزاوجة) Matching items

جـ- تجهيز واختيار الاجابة supply and select items وتسمى أيضاً الاجابة الحرّة وانتقاء الاجابة free response and choice- response على التوالي.

ومن الجدير ذكره أن الطالب في إجابته عن هذه الفقرات تكون مقيّدة، ورغم أن عددها كبير فإنه من اليسير تصحيحها بموضوعية كاملة بواسطة الحاسوب ما عدا فقرات (الاجابات القصيرة، والتكميل).

وتقدم في الصفحات القادمة نبذة عن كل نوع من الانواع المشار اليها أعلاه من الفقرات الموضوعية من حيث ميزاتها واستخداماتها والارشادات والقواعد الخاصة بتصميمها مع ذكر امثلة تطبيقية لكل نوع من تلك الأنواع.

الفقرات ذات الاجابات القصيرة Short answer items

وتسمى هذه الفقرات أحياناً، فقرات التكميل أو الاكمال أو الفراغات، وتصنف كفقرات موضوعية من نوع (تجهيز الاجابة Supply- type)، وتتكون من فراغ أو فراغين يقوم الطالب بملأها بكلمات، عبارات، أو رموز رياضية، أو معادلات، وغير ذلك.

وهناك ثلاثة أنواع شائعة الاستخدام من هذا الاختبار وهي:

أ- **نوع السؤال** Question حيث يوجه سؤال مباشر يتطلب من الطالب أن يجيب عليه بكلمة، جملة، عبارة والى غير ذلك.

مثال:
في أي مدينة تم انشاء اول مختبر لعلم النفس؟..
وعلى الطالب أن يكتب الاجابة في الفراغ بعد علامة الاستفهام، وهي (لايبزك).

ب- **نوع الاكمال**: Completion وفيه تقدم للطالب عبارة ناقصة، يكملها في الفراغ لتصبح عبارة كاملة صحيحة علمياً.

مثال:
إن المدينة التي تم فيها انشاء أول مختبر لعلم النفس هي.......
ويكتب الطالب كلمة (لايبزك) في الفراغ.

جـ- **نوع التداعي (الترابط)** Association وفيه تقدم للطالب كلمات أو مصطلحات معينة وعليه أن يكتب أمام كل كلمة ما يشابهها في المعنى أي (مرادفاً لها) أو ما يرتبط معا بعلاقة معينة.

مثال:

بعد إسم كل مدينة من المدن الآتية أكتب إسم الدولة التي تقع فيها.

- الموصل . . . .
- الدوحة . . . .
- صلالة . . . .

حيث على الطالب أن يملأ الفراغات بأسماء الدول وعلى التوالي: العراق، قطر، عمُان.

مثال: أكتب المرادف الصحيح للكلمات أو المصطلحات الآتية:

- يتقهقر . . . .
- يتلاشى . . . .
- يتذبذب . . . .
- يزدرد . . . .

ويرى العديد من الباحثين أن هذا النوع من الفقرات يقع بين الفقرات المقالية والفقرات الموضوعية. كما أن هذه الفقرات تتطلب من الطالب استدعاء او تذكر المعلومات، ويمكن تصحيحها بموضوعية إذا ما أحسن صياغتها؛ بحيث لا تتطلب إلا إجابة صحيحة واحدة فقط لا شك فيها.

استخدامات هذه الفقرات:

يمكن استخدام هذه الفقرات لتقويم مجالات متنوعة مثل: تقويم قدرة الطالب على تهجأة الكلمات كما في دروس اللغة Spelling كما إنها مفيدة للأستخدام في مواد الرياضيات والعلوم عندما يطلب من الطالب في عملية الجمع اعطاء النواتج لهذه العملية. كما يمكن في دروس العلوم استخدامها لتقويم قدرة الطالب على إكمال بعض رموز المعادلات الكيماوية مثلاً أو نواتجها. وقد وجد كل من فورسيث وسبرات (Forsyth and Sprott,1980) ان فقرات هذا النوع التي تتطلب عمليات حسابية تتفوق على فقرات الصواب والخطأ، وفقرات الاختيار من متعدد التي تتطلب تقويم القدرة على التعرف.

كما أن هذه الفقرات تستخدم أيضاً في تقويم معرفة الطالب للتعريفات والمصطلحات الفنية، وفي هذا المجال الذي يتطلب تجهيز الاستجابة Supply تكون هذه الفقرات أفضل من الأنواع الأخرى التي تعتمد على الاختيار Select.

أمثلة:

● لتقويم معرفة (الأدوات).

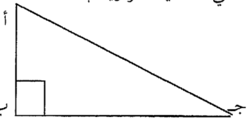
- أي أداة يمكن إستخدامها لقياس درجات الحرارة؟ الترمومتر .

● لتقويم معرفة حقائق محددة :

- ان مدة الدراسة في المرحلة الاولى للتعليم الأساسي في سلطنة عُمان أمدها . . . سنوات.

● لتقويم القدرة على تفسير البيانات :

- في الشكل التالي، ما قيمة الزاوية ب؟



● لتقويم قدرات التفكير العليا :

ما هو العدد الذي اذا أضيف لضعفه (5) يصبح (39)؟ . . . .

وأخيراً يمكن إجمال محددات استخدام الاجابات القصيرة كما يأتي:

1- نظراً لأنَّ فقرات الأسئلة القصيرة مناسبة لقياس قدرة الطالب على تذكر التواريخ، الأسماء، الاماكن أو المصطلحات فانها قد تشجع الطالب على الحفظ الآتي وعادات دراسية قاصرة.

2- تصحيح إجابات هذا النوع من الفقرات يستغرق وقتاً أطول منه في الانواع الموضوعية الاخرى بسبب قيام الطالب بكتابة جملة أو عبارة وبذلك تقترب من النوع المقالي، وقد لا يكون التصحيح دقيقاً بسبب إحتمال وجود أكثر من إجابة صحيحة لنفس السؤال وكما في المثال التالي:

أين تقع مدينة مسقط؟ ______ .

ويُلاحظ القارئ أن هناك عدة اجابات صحيحة لهذا السؤال القصير منها: في سلطنة عُمان، على بحر عُمان، جنوب شبه جزيرة العرب، وإلى غير ذلك. وسنأتي لاحقاً على القواعد والارشادات التي تحدد مثل هذا التنوع للاجابة الصحيحة.

3- إن هذه الاسئلة محددة باجابات على شكل كلمات، عبارات، رموز أو أرقام، وبذلك تكون الفرصة ضيقة لتقويم القدرة على، التعميم generalization أي تعميم الافكار والقواعد على حالات مشابهة، والتجريد abstraction أي عرض الفكرة بشكلها الحقيقي المجرد، والتفسير interpretation أي تفسير الظاهرة وتحليل العلاقات بين عناصرها، أو تبرير وجود الظاهرة لأن مثل هذه القدرات لا يمكن تقويمها لدى الطالب عن طريق كلمات قليلة محددة.

4- يصعب في هذا النوع تقويم قدرة الطالب على التركيب أو التوليف Synthesis لمحدودية الاجابةِ على هذا النوع واقتصارها على كلمة، أو رمز أو رقم.

إرشادات ومقترحات لكتابة فقرات جيدة من هذا النوع:

1- يجب صياغة الفقرة بدقة بحيث لا تتحمل غير إجابة صحيحة واحدة لا خلاف عليها. لا سيما في نوع الاكمال (الفراغات).

مثال قاصر:

- بغداد هي----.

ان هذه العبارة تتحمل أكثر من إجابة صحيحة واحدة مثل: (عاصمة العراق، مدينة جميلة، مركز العراق التجاري، أكبر من البصرةِ ... الخ).

وعليه فان هذه الفقرة يجب أن تعاد صياغتها كاللآتي:

- بغداد هي عاصمة ...... أو - بغداد أسسها الخليفة ......

2 - عند صياغة فقرات من نوع الأكمال يجب حذف الكلمة أو المصطلح المهم في العبارة من الناحية العلمية والذي يستهدف تقويم قدرة الطالب

على استرجاعه باعتباره أحد العناصر أو العنصر الاساسي في المفهوم، أو القانون، أو التعريف، وأن لا تستخدم هذا النوع لقياس القدرة على تذكر فعل، حرف جر، أو أحد أدوات العطف أو الربط وغير ذلك، إلا إذا كان الاختبار يتعلق باللغة العربية وقواعدها.

مثال قاصر: كولومبوس. (إكتشَفَ) أمريكا عام ....

بينما المفروض أن نسال الطالب عن اسم المكتشف لأمريكا، أو السنة التي اكتُشِفَت فيها تلك القارة:

مثال جيد: إكتشف كولومبس قارة أمريكا عام ..... ميلادية.

3 - يجب أن لا تتضمن الفقرة الواحدة العديد من الفراغات، لأن ذلك يؤدي الى ان يستخدم الطالب مجموعة من الخيارات كلها صحيحة مما يُفقد عملية التصحيح الاساسي الموضوعي من جهة، ومن جهة أخرى فإن ذلك يجعل من الصعب على الطالب التأكد من الموضوع أو المصطلح المراد وضعهُ في هذه الفراغات ليمثل الاجابة الصحيحة، فليس صحيحاً وضع الفقرة بالشكل الأتي:

مثال قاصر:

في معركة ..... التقى جيش .... مع جيش .... وكانت النتيجة اندحار جيش ....

ولو أمعنا النظر في هذه الفقرة لوجدنا أنفسنا أمام عدة تساؤلات منها:

أ- اي معركة اراد المعلم إختبار معلومات الطالب عنها؟

ب- في أي سنة او عهد نشبت تلك المعركة؟

جـ- بين أي جيشين حدثت المعركة؟

د- ما هو غرض المعلم من اختيار هذه الفقرة؟ هل يريد ان يعرف الطالب اسم المعركة؟ ام الجيشان اللذان دارت المعركة بينهما؟

هـ- أليس من حق الطالب أن يختار أي معركة ومن التاريخ القديم او الحديث وتكون اجابته صحيحةً؟

إن مثل هذه الصياغة لفقرات هذا الاختبار لا تشوش تفكير الطالب وحسب، بل أنها إضافةً الى ذلك تجعل بعض الطلبة لا سيما المتفوقين يتذمرون من إسلوب المعلم في إختياره لمثل هذه الفقرات التي لا تسعى لتحقيق أهداف مفيدة. وأخيراً لا بد من توجيه النصح للمعلم بتجنب الاكثار من عدد الفراغات في الفقرة الواحدة ويمكن تحديدها بفراغٍ واحد أو اثنين على الاكثر.

4 - ضرورة تحاشي نقل العبارات حرفياً من الكتاب المدرسي واستخدامها كفقرات في هذا الاختبار؛ لأن ذلك سيدفع الطالب الى حفظ المعلومات الموجودة في الكتاب حفظاً آلياً دون فهمها واستيعابها أو التدرب على تطبيقها، وهذا ما لا يستحب من الناحية العلمية.

5- يجب أن يكون الفراغ في نهاية الفقرة بدلاً من وسطها أو بدايتها، لأن وضع الفراغ في بداية الفقرة يحرم الطالب من فهم العناصر الاساسية لها ومن ثم التفكير في المصطلحات أو المعلومات المطلوب وضعها في الفراغ المحدد.

مثال قاصر: . . . . . .
. . . مادة ضرورية لحفظ توازن السكر في دم الإنسان.

مثال جيد: . . . . . .
- ان المادة الضرورية لحفظ توازن السكر في دم الانسان تسمى. . . (الانسولين).

6 - إجعل الفراغات الموجودة في جميع فقرات الاختبار متساوية في طولها، حتى لا يكتشف الطالب إن فراغٍ ما يحتاج الى كلمة واحدة وآخر لكلمتين وهكذا.

مثال قاصر:
إن المرض الذي يصيب الكبد لدى الافراد الذين يفرطون في تناول المشروبات الكحولية بدون تغذية كافية . . . .

ان وجود الفراغ بهذا الطول الواضح قد يدفع الطالب الى التخمين بأن الاجابة الصحيحة تتطلب كتابة أكثر من كلمة في هذا الفراغ.

7 - إذا كانت الفقرة تتألف من سؤال قصير أو عبارة ناقصة تتطلب إجابة رقمية؛ فيجب تحديد إسم الوحدة المطلوبة بعد الفراغ مباشرةً لأن ذلك من شأنه ان يُسِّط مشكلة التصحيح ويزيل اي مصدر للغموض قد يجعل المعلم في موقف صعب في تقرير ما إذا كان الطالب يعرف الاجابة حقيقةً أم لا.

مثال قاصر:

- ما كمية الحليب التي يوصى بها كحد أقصى لما يتطلبه فتىً في الرابعة عشرةَ من عمرهِ؟ فقد تكون الإجابة: 4 أكواب، لتر واحد، 1000 سنتمتر مكعب.

والأفضل أن تُصاغ الفقرة بالشكل التالي:

- ان كمية الحليب التي يُوصى بها كحدٍ أدنى لما يتطلبهُ فتىً في الرابعة عشرة من عمره هي .... أكواب.

إختبار الصواب والخطأ True- false

وهو واحد من الاختبارات التحصيلية الموضوعية الشائعة الاستخدام في الميدان التربوي بسبب سهولة تصميمه لا سيما عندما يركز على قياس الجانب المعرفي لدى الطالب، وكذلك سهولة تصحيحه يدوياً أو آليا. وتهدف فقرات هذا النوع الى تقويم قدرة الطالب على الدقة والتمييز بين الفقرات الصحيحة وغير الصحيحة من الناحية العلمية عندما تعرض عليه. أما من حيث شكل الفقرات فهي ثنائية الإجابة، وتكتب بأشكال عديدة وكالآتي:

1- (صح) (خطأ) يعتبر نهر النيل من أطول الانهار في العالم.

2- (صح) (خطأ) تزداد سرعة الضوء كلما إزداد الضغط الجوي.

3- تحدث ظاهرة الخسوف عندما تقع الأرض بين الشمس والقمر. ( ).

4- وهناك شكل آخر يعتبر من أفضل أنواع صياغة واستخدام فقرات الصواب والخطأ وفيه تُعرض على الطالب خرائط أو رسوم بيانية، أو قصة قصيرة

تعقبها فقرات تتعلق إجاباتها بتلك المقدمات بشكل مباشر أو غير مباشر، كما سنوضح ذلك لاحقاً.

5- وهناك شكل آخر يطلب فيه من الطالب أن يقرأ الفقرة ويقرر ما اذا كانت صحيحة أو خاطئة اولاً، فاذا أجاب ان الفقرة خاطئة عليه ان يبين السبب في خطئها، أو أن يصحح الخطأ لتكون الفقرة بشكلها الصحيح.

مثال :
- ان مصطلح " اللاشعور " أكدت عليه المدرسة السلوكية في علم النفس.
صواب.   خطأ.

فاذا ما اختار الطالب كلمة (صواب) فسيحصل على (صفر) لأن اجابته خاطئة تماماً. أما اذا اختار كلمة (خطأ) فعليه ان يبين الخطأ ويصححه، وفي هذه الحالة عليه أن يحدد السبب هو أن المدرسة هي (مدرسة التحليل النفسي) وليس (المدرسة السلوكية) وعندها تكون اجابته صحيحة تماماً.

ويُفسر البعض السبب في إتباع هذا الاسلوب مع إختبار الصواب والخطأ هو اختبار قدرة الطالب وثقته من دقة الإجابة من جهة، وتقليل استخدام التخمين اي الاجابة بالصدفة. لكن بعض الباحثين الآخرين يرون ان استخدام هذا الاسلوب يخرج فقرات الصواب والخطأ من النوع الموضوعي واقترابه من الاختبارات المقالية لان الطالب سيكتب عبارة أو عدة عبارات كما يفعل في حالة الاختبارات المقالية. كما يطرح البعض الآخر كيفية توزيع العلامة المخصصة للفقرة بين الاشارة التي يختارها، والتفسير الذي يقدمه.

إستخدامات إختبار الصواب والخطأ:

بشكل عام يستخدم معظم المعلمين هذا النوع لقياس الجانب المعرفي عند الطالب، لذلك يبدو هذا الاختبار للكثيرين بأنه اختبار بسيط لا يتعدى السؤال عن اسماء مدن أو تواريخ أو علماء أو مخترعات وغير ذلك.

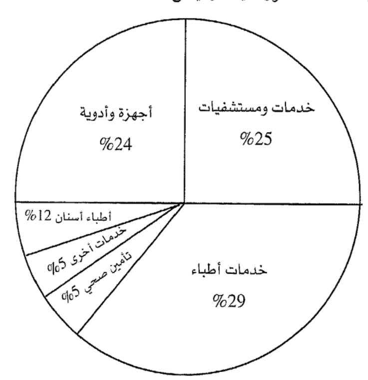
وفي رأينا أن المعلم اذا أراد ان يمزج بين الانواع الموضوعية والمقالية في

اختبارات طلابه فبالامكان استخدام هذا النوع على أن يحدد لطلابه ذلك ويوضح لهم أسلوب توزيع العلامة على المهارتين (الاشارة والتفسير والتصحيح).

لكنه لو يُصاغ بشكل جيد من قبل معلم ذو خبرة وتمرس في مجال صياغة الأختبارات لأصبح مناسباً لقياس قدرات أخرى اكثر تعقيداً من مستوى المعرفة، كالفهم والاستنتاج والتفكير المنطقي لا سيما اذا ما قُدِّمَت للطالب مقالة أو تجربة علمية مختبرية تعقبها فقرات من الصواب والخطأ.حيث تكون هذه المقدمات بمثابة اطار مرجعي يستخدمها الطالب لاختيار الاجابة الصحيحة وترك الاجابة الخاطئة.

مثال (1)

يتضمن الرسم الدائري المدرج في أدناه ستة قطاعات تمثل أوجه الإنفاق في مجال الرعاية الطبية المختلفة بالدولارات في الولايات المتحدة المطلوب منك تدقيق النظر في ذلك الشكل قبل أن تقرأ العبارات التي تعقبهُ لتقرر أي منها صحيحة لترسم دائرة حول الحرف (ص)، أو خاطئة لترسم دائرة حول الحرف (خ) . . . . . (ثورندايك وهيجن، 1989).



العبارات:

1- أن ما يُنفق في قطاع (خدمات الاطباء) أكثر مما ينفق في القطاعات الاخرى مجتمعةً ص، خ.

2- هناك قلة في عدد الموظفين الحاصلين على التأمين الصحي. ص.خ.

3- أن ما خصص للأجهزة الطبية والأدوية يمثل ثلث المجموع العام للإنفاق. ص.خ.

4- أن ما خصص لطب الأسنان لا يوازي ما خصص لخدمات المستشفيات. ص. خ

5- ان ما ينفق على التأمين الصحي والاشكال الاخرى للرعاية الطبية أقل مما يُنفق على طب الاسنان. ص.خ.

مثال (2)

في التجربة الموضحة في أدناه هناك إنبوبتي اختبار (س)و(ص) كل منهما مقعرة من جانبها. مُلأت الانبوبتان (بماء الجير) حتى الثلث ووضعت فوق مكان التقعر بكلا الانبوبتين شبكة سلكية، ووضعت فوق الشبكة السلكية في الانبوبة (س) فقط حشرة الخنفساء، بينما تركت الثانية بدون اي حشرة بداخلها. أُغلقت فوهتا الانبوبتان بسداد قطني وتُركتا لمدة (20 دقيقة)، وبعد انتهاء المدة المذكورة لوحظ تعكر ماء الجير في الانبوبة (س) لكنه ظلّ صافياً في الأنبوبة (ص) (Lindvall, 1968)

إن هذه التجربة تقيم الدليل على ان:

(ص)(خ) 1- ثاني اوكسيد الكربون كان يطلق حراً في القنينة (س).

(ص)(خ) 2- الخنفساء لا تستطيع العيش بالقرب من ماء الجير.

(ص)(خ) 3- الاوكسجين كان يستنفذ في الانبوبة (س).
(ص)(خ) 4- الأوكسجين كان يطلق حراً في الانبوبة (س).
(ص)(خ) 5- ثاني اوكسيد الكربون كان يستنفذ في الانبوبة (ص).

ومن الفوائد والاستخدامات الاخرى لهذا النوع من الاختبارات ما يأتي:

1- في هذا النوع من الاختبارات يستطيع المعلم أن يغطي مساحة واسعة من المادة المقررة، أي إنها تحقق مبدأ الشمول بشكل أكبر.
2- سهولة اجابة الطالب على هذا النوع من الفقرات وقلة الوقت المستخدم في الاجابة مقارنةً بالانواع الاخرى من الفقرات الموضوعية، وقد وجَدَ بعض الباحثين أن الطالب يجيب على ثلاث فقرات من هذا النوع مقابل فقرتين من نوع الاختيار من متعدد في نفس الوقت .(Frisbie.1974)
3- يُناسب القدرات العقلية للطلاب بكافة فئاتهم العمرية ابتداءً من الصغار في المرحلة الابتدائية صعوداً الى المرحلتين الثانوية والجامعية.

عيوب إختبار الصواب والخطأ

لقد سجل بعض الباحثين من خلال التطبيق الميداني لفقرات هذا الاختبار العديد من عيوب ونقاط ضعف، ومنها:

1- التخمين: حيث تلعب الصدفة دوراً كبيراً يصل الى (50%) في الاختيار الصحيح للاجابة دون الاعتماد على الفهم والتأكد من الاجابة.
2- سهولة الغش في هذا النوع، فنظرة سريعة للطالب بورقة زميله المجاور له في قاعة الامتحان يستطيع التقاط عدة اجابات مرةً واحدة وبسهولة من تلك الورقة.
3- قد تشجع الطالب على الحفظ الآلي، لا سيما اذا ما اقتبس المعلم عبارات بالنص من الكتاب المقرر.
4- ان سهولة الكثير من فقرات هذا الاختبار تؤدي الى نجاح معظم طلبة الفصل، مما ينتج عن ذلك إنخفاض معاملات تمييز هذه الفقرات، ويقلل من فائدة استخدامها في تحديد الطلبة الاقوياء والطلبة الضعفاء.

5- ان معامل الثبات لهذا النوع من الفقرات أقل منه في أنواع اخرى من الفقرات الموضوعية، فقد وجد كل من فرسبي (Frisbe,1974) وايبل (Ebel, 1975) ان معامل ثبات الصواب والخطأ أقل من معامل ثبات الاختيار من متعدد.

إرشادات وقواعد لصياغة فقرات الصواب والخطأ:

1- يجب إختبار عبارات مجددة، أي إما صحيحة تماماً او خاطئة تماماً ولا يوجد شك في صحتها أو خطأها. فليس صحيحاً طرح العبارة التالية:

مثال:

ص. خ. البنسلين علاج فعال في معالجة أمراض الانسان.

ومن الواضح أن هذه العبارة تتحمل اجابتها اختيار (الصواب او الخطأ) فقد يعتقد بعض الطلبة أن هذا العلاج يشفي أمراض الانسان بينما الطالب الذكي المطّلع على معلومات صحية أكثر، قد يعتبر الاجابة الصحيحة هي اختيار الحرف (خ) لانه يعرف أن البنسلين قد يُشفي بعض الامراض عند الانسان وليس جميعها، وهنا يواجه المعلم حالة ارباك في اعتبار من اختار (ص) هو الصحيح ام من اختار (خ) هو الصحيح لان كلا الاختيارين صحيحاً. والافضل ان يحدد المعلم في صياغة العبارة اتجاه واحد قاطع للاجابة. أي أن تعاد صياغة المثال ليصبح كالآتي:

ص.خ. البنسلين دواء فعال لمعالجة مرض التدرن الرئوي.

2- تجنب اختيار العبارات المركبة أي التي تتضمن أكثر من فكرة واحدة لا سيما التي تتضمن فكرتين احداهما صحيحة والاخرى خاطئة؛ لأن ذلك يربك معظم الطلاب في كيفية تحديد الاجابة الصحيحة. فليس صحيحا طرح العبارة بالشكل الآتي:

مثال:

ص.خ. يقع العراق في قارة آسيا، جنوب الكويت.

ومن خلال استقراء المثال اعلاه نلاحظ أن النصف الاول المتمثل في موقع العراق بالنسبة لقارة آسيا صحيحاً، بينما النصف الثاني (جنوب الكويت)

خاطئاً. وفي هذه الحالة يمكن أن يقسم المعلم العبارة الى عبارتين مستقلتين. ولا بد من الاشارة هنا الى أن بالامكان استخدام مثل هذه العبارة المركبة اذا كان السؤال يتطلب اكتشاف الخطأ وتصحيحه في مجموعة عبارات تطرح على الطالب.

3- تجنب استخدام العبارات المصاغة بالنفي أو نفي النفي، اذ لوحظ أن الكثير من الطلبة يغفل اشارة النفي عند اجابته على هذا النوع من الفقرات، أما نفي النفي فهو مُعقد وكثيراً ما يُثير الالتباس عند الطالب.

مثال خاطىء:

ص. خ. لا يقوم الكلوروفيل بالتركيب الضوئي في النبات الا اذا لم تتوفر أشعة الشمس.

وفي هذا المثال إن العبارة منفية باشارة (لا) في بدايتها، وهناك علامة نفي النفي (لم) في الجزء الثاني من العبارة.

مثال جيد:

ص.خ. يقوم الكلوروفيل بالتركيب الضوئي في النبات في حالة عدم توفر أشعة الشمس. مع ملاحظة ان على المعلم اذا ما اضطر ان يستخدم علامة النفي أن يضع خط صغير تحتها ليلفت انتباه الطالب على ان الجملة منفية.

4- على مصمم هذا النوع من الفقرات تجنب استخدام كلمات (التعميم والتخصيص) التي يُشير إستخدامها الى اختيار الاجابة الصحيحة لاقتران تلك المصطلحات بها، فالمصطلحات مثل (كل، جميع، قطعاً، دائماً، حتماً) يدرك منها الطالب انها لا تنطبق على حالات استثنائية بل تتميز بالعمومية مما يستنتج منها الطالب خطأ العبارة المستخدمة ضمنها.

مثال خاطىء:

It is always better to use longer test than a short one. T.F

أما العبارات التي تتضمن صيغاً او مصطلحات فيها نوع من الأشتراط أو الاستثناء من مثل (في بعض الأحيان، ربما، تحت ظروف معينة... الخ) فغالباً ما تكون العبارات التي تتضمنها صحيحة، ويمكن للطالب المتفوق استنتاج صحة أو خطأ العبارات من خلال هذه المصطلحات.

5- من الضروري ان تكون فقرات الصواب والخطأ قصيرة ومختصرة كلما أمكن ذلك، وان تتجنب حشو الفقرة بعبارات تزيد من طولها ولا حاجة لها كالخوض بالتفاصيل التي لا علاقة لها بالموضوع المراد اختبار الطالب فيه كما في المثال الآتي:

مثال خاطىء:

ص. خ. أن جمهورية السودان التي تحدها جمهورية مصر من الشمال وينبع منها نهر النيل وهي أحد أعضاء جامعة الدول العربية تقع في قارة أسيا.

ونلاحظ في المثال اعلاه ان الهدف من الفقرة إختبار قدرة الطالب على تحديد اسم القارة التي تقع فيها جمهورية السودان، لكن واضح ان الفقرة أضيفت إليها معلومات كثيرة لا علاقة لها بهذا الهدف، متناسياً أن هذا الإطناب في الكتابة استغرق منه وقتاً اطول في كتابة الفقرة وفي طباعتها، واحتل من ورقة الاختبار مساحة أكبر، إضافةً إلى أنه سيرهق الطالب في القراءة والتفكير والنظر في مدى علاقة هذه المعلومات بالهدف الاساسي من الفقرة وكل هذه الامور مضيعات للوقت والجهد لكلٍ من المعلم والطالب.

6- يفضل أن يكون عدد الفقرات الخاطئة مساوياً لعدد الفقرات الصحيحة، وأن يكون طولهما متساوياً، وأن يوزعا على الاختبار ترتيب عشوائي، أي غير مقصود حتى لا يكشفه الطالب فيستخدم التخمّين بدلاً من التفكير.

7- في الفقرات التي تصف العلاقة بين متغيرين أحدهما سبباً والآخر نتيجة، يجب ان لا نكتب السبب أولاً والنتيجة ثانياً بحيث يستنتج الطالب صحة العلاقة من خلال الترتيب، وكما في المثال التالي:

> **مثال خاطىء:**
> ص. خ. أن تفاعل الكلور مع الصوديوم ينتج عنه تكون ملح الطعام. والمثال الصحيح هو: ص.خ. ينشأ ملح الطعام من تفاعل الصوديوم مع الكلور.

8 - لا يفضل إقتباس عبارات بشكل حرفي من الكتاب المقرر ووضعها كفقرات صواب وخطأ، لان ذلك قد يؤدي إلى تعلم الطالب لظاهرة الحفظ الآلي، وعلى مصمم الاختبار ان يأخذ الفكرة من الكتاب المقرر، ويعيد صياغتها بأسلوبه الخاص، ويضعها ضمن فقرات هذا النوع من الاختبارات.

9- تجنب ذكر آراء لم يتم الاتفاق عليها بشكل نهائي من الناحية العلمية فمثلاً ليس صحيحاً ذكر الفقرات الآتية:

> **مثال :**
> ص. خ. ان التدخين يؤدي إلى الإصابة بمرض السرطان.
> ص. خ. ان افكار فرويد عن الغريزة الجنسية مطابق للحقيقة والواقع.

10- وأخيراً لا بد لنا أن نُذّكر المعلم أو من يصمم هذا النوع من الفقرات الى أمرين ضروريين. الاول يتعلق بكيفية معالجة مشكلة (التخمين guessing)

حيث يطرح بعض الباحثين بامكان المعلم تحذير طلابه بضرورة عدم استخدام هذه الظاهرة وضرورة ترك الفقرة التي لا يعرف الطالب اجابتها بشكل صحيح حتى لا تحسب عليه عند استخدامه لمعادلة معالجة أثر التخمين أثناء تصحيح أوراق الاجابة بعد أن يبين لهم هذه المعادلة وتأثيرها على تحديد الدرجة النهائية للطالب قبل اجراء الاختبار وهي بشكلها البسيط التالي:

د= ص - خ

أي: الدرجة النهائية = مجموع الفقرات التي اجاب عنها الطالب إجابة صحيحة مطروحاً منها عدد الاجابات التي اخطأ فيها.

فمثلاً لو أن الطالب أجاب على 40 فقرة بشكل صحيح من أصل 50 فقرة، اي انه أخطأ بعشرة فقرات، فان درجته النهائية ستكون: د = 40 - 10 وتساوي (30).

ولعل الافتراض الذي يقع خلف هذه المعادلة في حذف عدداً من الدرجات مساوياً لعدد اجاباته الخاطئة، على اساس ان هذا الطالب اجاب بشكل صحيح عن (30) فقرة بدون تخمين، وبقيت عنده (20) فقرة لم يستطع الاجابة عنها فاستخدم عملية التخمين افتراضاً. أما الأمر الثاني الذي يجب الانتباه له فهو موضوع (الغش cheatting) الذي يسهل في هذا النوع من الفقرات، فان بالامكان معالجته في طبع الفقرات بترتيبات مختلفة لكل صورة (form) من صور الاختبار وتوزيعها بشكل مختلف على الخطوط التي يجلس فيها الطلبة في قاعة الاختبار.

إختبار الإختيار من متعدد Multiple choice

يعتبر اختبار الاختيار من متعدد من أكثر الاختبارات الموضوعية انتشاراً في مجال القياس والتقويم سيما وانه يقيس مخرجات تعليمية متنوعة من أبسط مستوياتها الى أكثرها تعقيداً ويمكن صياغة فقراته بنوعية عالية (Popham, 2006) يتكون هذا الاختبار من جزئين، الأول يسمى أصل السؤال أو (البند) stem الذي يمثل مشكلة، اما الجزء الثاني فهو عبارة عن بدائل alternatives أو اختيارات options or choices التي تمثل حلول محتملة للمشكلة أو الاصل وتتضمن الاجابة الصحيحة واجابات اخرى خاطئة تسمى المموهات distrac-ters ووجودها يمثل خداعاً للطلبة غير المتأكدين من الاجابة. ان أصل السؤال أو البند يمكن صياغته باشكالٍ متعددة كالآتي:

أولاً: صيغة السؤال القصير (الاستفهام):

مثال(1):
أي من الاختبارات الموضوعية التالية أكثر ملائمةً لقياس معظم الأهداف السلوكية؟
أ- الاختيار من متعدد.
ب- الصواب والخطأ.
جـ- المزاوجة.
د- الإجابات القصيرة.

مثال (2):
- أي من نظريات علم النفس الآتية تؤكد دور التعزيز في عملية التعلم؟
أ- السلوكية. ب- المعرفية.
جـ- الاجتماعية. د- التحليل النفسي.

ثانياً: قد يكون أصل السؤال عبارة ناقصة تكملة إحدى البدائل التي تعقب العبارة.

مثال (1):
- ان القطر العربي الذي فيه أكبر عدد من النخيل في العالم هو.
أ- السعودية. ب- عُمان.
جـ- العراق. د- مصر.

مثال (2):
- أن اول معركة خاضها المسلمون ضد المشركين هي:
أ- أحد. ب- الخندق.
ج- تبوك. د- اليرموك.

ثالثاً: قد يكون أصل السؤال كلمة واحدة ويطلب من الطالب أن يختار مرادفاً لها، أو تعريفا لها، أو ترجمةً لها من بين البدائل التي تعقبها.

مثال (1):

- يتذبذب:

أ- يفصل. ب- يهتز.

جـ- يُغير. د- يقبَل.

مثال (2):

- الخسوف:

أ- توسط الارض بين الشمس والقمر.

ب- وقوع القمر بين الارض والشمس.

جـ- توسط الشمس بين الارض والقمر.

د - وقوع الشمس والقمر والأرض على خط مستقيم واحد.

أما بالنسبة الى البدائل أو الاختيارات التي تعقب أصل السؤال فهي تمثل اجابات محتملة لاصل السؤال أو تكملات محتملة للعبارة الناقصة في أصل السؤال، أو مرادفات أو تعريفات للمصطلح او للكلمة الموجودة في أصل السؤال.

وهناك بعض الملاحظات المتعلقة بالبدائل نوجزها بالآتي:

1- في غالب الاحيان يكون عدد البدائل من ثلاثة حتى خمسة، ويتحدد العدد وفقاً للعمر العقلي والزمني للطالب.

2- يتعين أن تكون واحدة من الخيارات صحيحة والباقيات خطأ وتسمى الخيارات الخطأ بالموهات .disracters

3- في بعض اشكال هذا النوع يطلب من الممتحن أن يختار أفضل إجابة من بين الخيارات المعروضة عليه وجميعها تمثل اجابات صحيحة، ويستخدم هذا النوع لقياس التحصيل المعقد للطالب مثال ذلك عندما يطلب من الممتحن إختيار أفضل سبب لظاهرةٍ أو حدثٌ ما أو أفضل وسيلة لعمل شيءْ ما أو أفضل تطبيق لمبدأ ما وغيرٍ ذلك من الامور التي تعتمد على نوع الاداء المراد قياسه.

مثال :

أي من الاختبارات التالية يُعتبر الافضل لقياس المهارات الحسابية؟

أ- الاختيار من متعدد.

ب- الصواب والخطأ.

جـ- المطابقة.

د- الاجابات القصيرة.

4 - قد يطلب من الممتحن ان يحدد البديل المختلف عن البدائل الاخرى أو الذي لا تنطبق عليه مواصفات مذكورة في أصل السؤال، وفي هذا النوع يوجه انتباه الطالب نحو إختيار الاجابة الخطأ.

مثال :

أي من إناث الحيوانات التالية لا تلد صغاراً؟

أ- الحوت. ب- النسر.

جـ- التمساح. د- الزرافة.

5 - في بعض الاحيان يلجأ مصمم هذا النوع من الفقرات الى صياغة الاختيار (البديل) الاخير بكلمات شائعة مثل: كل ما ذُكرَ صحيحاً، الاجابة الصحيحة غير مذكورة . . . . الخ ,وهناك اتجاهين حول هذه البدائل، الاتجاه الاول: لا يؤيد استخدامها على أساس انها مربكة للطلبة ، بينما الاتجاه الاخر يؤيد استخدامها على اساس انها تقيس ثقة الطالب بمعرفته للمعلومة الخاصة بالفقرة. وقد يستخدم بعض المعلمين هذه الكلمات عندما لا يجدون بدائل أُخرى تتماشى مع البدائل المتعلقة بالموضوع الذي تدور حوله الفقرة.

ميزات الإختيار من متعدد

بالإضافة الى شيوع استخدام فقرات هذا النوع فانه يمتاز بمميزات ايجابية عديدة منها:

1- تقيس فقرات هذا النوع جميع أصناف الاهداف السلوكية من أبسطها حتى أعقدها، كما يمكن صياغتها من محتوى معظم المواد والمراحل الدراسية (Mehrens & Lehmann, 1991)

2- يمكن تقدير علامات هذا النوع بموضوعية كاملة مهما اختلف عدد المصححين، اضافةً الى ان عملية تصحيحه تستغرق وقتاً ليس كبيراً سواء باليد أو بالحاسب الآلي.

3- يتمتع هذا النوع بدرجة جيدة من الصدق بسبب شمول فقراته لمساحة أكبر من المحتوى الذي يُغطيه الاختبار مما يزيد من تمثيل ذلك المحتوى بشكل أفضل ويحقق الصدق الظاهري وصدق المحتوى في الوقت نفسه.
كما انه يتمتع بدرجة مقبولة من الثبات بسبب تمتعه بالصدق من جهة، ولأن كثرة عدد فقراته يؤدي الى زيادة طول الاختبار وما ينشأ عن ذلك من زيادة درجة ثباته.

4- إن مجال الحزر أو التخمين أقل في هذا النوع من الانواع الموضوعية الاخرى ويمكن تقليص نسبته بزيادة عدد البدائل. فعندما يكون عدد البدائل ثلاثة تكون نسبة التخيمن (33.3%)، وعند يصبح عدد البدائل أربعة تصبح النسبة (25%)، وتصبح النسبة (20%) عندما يكون عدد البدائل خمسة.

5- باستخدام هذا النوع يمكن تحقيق مستوىً من الشمول أكبر بسبب كثرة الفقرات والبدائل التي يتكون منها والتي تتوزع على أكبر عدد من موضوعات المحتوى.

6- نظراً للميزات الايجابية التي يمتاز بها هذا النوع نلاحظ أن معظم الاختبارات المعيارية التي تصدرها مؤسسات نشر الاختبارات في العالم هي من نوع الاختيار من متعدد.

عيوب هذا النوع:

على الرغم من الايجابيات التي يمتاز بها هذا الاختبار، فان هناك بعض العيوب التي ترافق استخدامه ومنها:

1- يحتاج اعداده وتصميمه الى وقت ودراية ومهارة أكبر من قبل المعلم مقارنةً بالانواع الاخرى من الاختبارات.
2- تحتاج إجابة هذا النوع الى وقت وجهد اكبر من الطالب نظراً لانه يقرأ اصل السؤال حتى يفهمه أولاً ثم يقرأ البدائل عدة مرات ليقرر اختيار البديل الصحيح.
3- يشغل هذا الاختبار حيزاً كبيراً من ورقة الاختبار بالاضافة إلى الحاجة لإمكانات طباعية قد لا تتوفر لدى بعض المعلمين.

قواعد وارشادات لتصميم فقرات الاختيار من متعدد

يمكن تقديم نوعين من القواعد والارشادات الخاصة بتصميم فقرات هذا الاختبار وهي:

أولاً: القواعد الخاصة بأصل السؤال:

1- صمم كل فقرة من فقرات الاختبار لتُقيس مخرجات تعلم مهمةً. ان الموقف الذي تدور حوله مشكلة ما لتصمم منه فقرة إختبارية يجب ان يكون مرتبطاً بمخرجة تعليمية نُريد قياسها. فالفقرة التي تصمم لقياس تحصيل معقد يجب أن تتضمن نوعاً من الجدة أو الحداثة، فلو فرضنا أن فقرة ما إستهدفت تحديد الطالب لتعريف مفهوم موجود في الكتاب المقرر، وفقرة اخرى تقيس الفهم تتطلب قدرة الطالب على تحويل ذلك المفهوم الى شكل آخر، وفقرة اخرى تقيس التطبيق تتطلب استخدام مصطلح معين في مجال معين. اذن كلا فقرتي الفهم والتطبيق جاءت وفق ما يقصده مصمم الاختبار وهذا يمكن ان يتحقق في حالة المادة الدراسية الجديدة للطالب.
2- يجب أن يُصاغ أصل السؤال بشكل مشكلة واضحة ومحددة يستطيع الطالب أن يفهمها حتى دون أن يقرأ البدائل اللاحقة. ويمكن ان يتحقق ذلك بشكل أفضل عندما نستخدم الصيغة الاستفهامية لأصل السؤال. ويمكن ملاحظة ذلك من خلال الامثلة الآتية:

صياغة ضعيفة:

1- النفط:

أ- سائل. ب- صلب. جـ- غاز؟.

2- جدول المواصفات:

أ- يُشير الى كيفية استخدام الاختبار لتحسين عملية التعلم.

ب- يوفر توازن نسبي لمحتوى المادة الدراسية.

جـ- يرتب الاهداف التدريسية حسب أهميتها.

د- يوضح كيفية استخدام اسلوب تصحيح أوراق الامتحان.

صياغة جيدة:

1- بأي من الحالات التالية يوجد النفط؟

أ- السائلة. ب- الصلبة. جـ- الغازية.

2- ما هي الفوائد المتوخاة من استخدام جدول المواصفات عند تصميم الاختبار؟

أ- يقلل الوقت المستخدم في بناء الاختبار.

ب- يوفر توازن نسبي لمحتوى المادة الدراسية.

جـ- يجعل عملية بناء الفقرات أيسر.

د- يزيد من موضوعية الاختبار.

ويلاحظ من المثالين السابقين أن الصياغات الاولى لهما لا تشجع الطالب على التفكير كما يلاحظ من الصياغات الجيدة، حيث يقرأ الطالب أصل السؤال (المشكلة) ثم يستعرض الحلول (البدائل) ليختار أفضلها.

3- يُفترض صياغة أصل السؤال بلغة بسيطة واضحة، والابتعاد عن استخدام الكلمات المعقّدة، والاختصار في الصياغة وعدم حشو السؤال بكلمات غير ضرورية، لأن وجود معلومات كثيرة غير ضرورية تستنفذ من الطالب وقتاً أكبر وقد تعوق تفكيره في فهم أصل السؤال ومن ثم اختيار البديل الذي يمثل الاجابة الصحيحة.

مثال ضعيف:

ينبع نهرا دجلة والفرات من الأراضي التركية، يمر الفرات عبر الاراضي السورية قبل دخوله العراق، أما دجلة فيدخل الاراضي العراقية من الشمال بعدَ الاراضي التركية مباشرةً، وفي العراق تتفرع منهما فروع كثيرة، ما إسم المنطقة التي يلتقي فيها النهران في العراق؟

أ- سوق الشيوخ، ب- الغراف، جـ- الكحلاء، د- القرنة.

مثال جيد:

ما اسم المدينة التي يلتقي عندها نهرا دجلة والفرات في الاراضي العراقية؟

أ. سوق الشيوخ، ب- الغراف، جـ- الكحلاء، د- القرنة.

4- يجب وضع المصطلحات الضرورية في اصل السؤال بدلاً من اعادتها مع كل بديل من البدائل المعطاة، لأن ذلك يستنفذ وقتاً وجهداً اكبر من الطالب الذي يضطر الى اعادة قراءة نفس المصطلح عدة مرات بحكم وجوده مع كل بديل.

مثال ضعيف:

- ان مصطلح (الموضوعية) في الاختبار الموضوعي:

أ- يُشير الى اسلوب تحديد الناتج التعليمي.

ب- يُشير الى اسلوب اختيار محتوى الاختبار.

جـ- يُشير الى اسلوب عرض المشكلة.

د- يُشير الى اسلوب تصحيح الاجابات.

مثال جيد:

الصياغة الأفضل لهذه الفقرة يمكن أن تكون بالشكل الآتي:

- إن مصطلح (الموضوعية) في الاختبار الموضوعي يُشير إلى اسلوب:

أ- تحديد الناتج التعليمي.　　ب- إختيار محتوى الاختبار.

جـ- عرض المشكلة.　　د- تصحيح الاجابات.

5 - ضرورة صياغة أصل السؤال بصيغة الاثبات وتجنب الصياغة بالنفي حيث لوحظ ان بعض الطلبة لا ينتبهون الى إشارة النفي، واذا ما اضطر المعلم او مصمم الاختبار لإستخدام إشارة النفي فيجب وضع خط صغير تحتها، أو طبعها باللون الغامق Bold حتى ينتبه الطالب الى ان هذه العبارة منفية فيأخذ ذلك بنظر الاعتبار.

مثال :

- أي مما يأتي لا ينتمي الى مقاييس النزعة المركزية؟

أ- الوسط. ب- الوسيط. جـ- المدى د- المنوال.

إن ما تقدم ذكره لا يمنع من استخدام بعض الفقرات الاختبارية بالصيغة السالبة، لان هناك بعض المعلومات او الحقائق تكون الصياغة السالبة لها أمر ضروري لذلك تصبح جزءاً من عملية القياس، فمثلاً يستخدم المعلم في تدريسه كلمات تحذيرية أو ارشادية ذات صياغة سالبة منها: لا تعبر الشارع والاشارة الضوئية حمراء، أو لا تلعب بمواد كيمياوية لا تعرف طبيعتها، وهكذا نرى ان العديد من الحقائق تتطلب اسلوب النفي لذلك لا يمنع من الصياغة السالبة في مثل هذه الظروف. ويمكن اعادة صياغة المثال السابق بإستبدال كلمة (لا) بكلمة (ما عدا) ووضعها في نهاية أصل السؤال لتكون واضحة أكثر للطالب أثناء قراءتها ويمكن طبعها بلون غامق وكالآتي:

مثال :

- جميع الأساليب الاحصائية الآتية من مقاييس النزعة المركزية ما عدا:

أ- الوسط. ب- الوسيط. جـ- المدى د- المنوال.

6- يُفترض صياغة أصل السؤال بدقة بحيث لا يضطر الممتحن الى التكهن بما يقصده واضع الاختبار، أي أن يحتوي على جميع المعلومات اللازمة، وان بدت بديهية.

مثال ضعيف :

- يغلي الماء بدرجة حرارة:

أ- (50م)   ب- (100م)   جـ- (150م)   د-(٢٠٠م)

ويلاحظ أن المثال أعلاه غير دقيق علمياً وتنقصه حقائق علمية اخرى لذلك يجب ان تعاد صياغته بالشكل الآتي:

- يغلي الماء النقي عند مستوى سطح البحر وثبوت الضغط الجوي بدرجة حرارة:

أ- (50م)   ب- (100م)   جـ- (150م)   د- (200م).

7- اذا أراد المعلم أو مصمم الاختبار تقويم قدرة الطالب على معرفة او فهم مصطلح علمي معين، فمن الافضل أن يضع تعريف ذلك المصطلح في أصل السؤال، بينما يضع المصطلح المقصود مع مصطلحات اخرى كبدائل ولا يفضل العكس حتى يوفر الوقت والجهد على الطالب الذي عادةً ما يقرأ الأصل مرة واحدة بينما يحتاج الى قراءة البدائل أكثر من مرة ليقرر اختياره الصحيح. لكننا في بعض الأحيان قد نواجه صعوبة الحصول على عددٍ من المصطلحات المنسجمة لوضعها كبدائل للتعريف.

مثال :

- تسمى ظاهرة توسط الأرض بين الشمس والقمر ظاهرة:

أ- الخسوف.   ب- الكسوف.   جـ- ـــــــ.   د-ـــــــ.

وهنا لا يُوجد أكثر من مصطلحين لوضعهما في قائمة البدائل، لذلك يمكن تقديم ثلاثة بدائل يختار المعلم واحداً منها وهي:

أولاً: وضع مصطلح (الخسوف) في أصل السؤال والتعريفات كبدائل وكما ياتي:

مثال :
- يحدث الخسوف نتيجةً:
أ- توسط القمر بين الشمس والارض.
ب- وقوع الارض بين الشمس والقمر.
جـ- وجود الشمس بين الارض والقمر.

ثانياً: تحويل المثال الى فقرة (صواب وخطأ)، وكما يأتي:

يحدث ظاهرة الخسوف نتيجة توسط الارض بين الشمس والقمر. (ص)،(خ).

ثالثاً: أما البديل الثالث فهو الاستغناء عن تصميم فقرة حول هذا الموضوع.

8- يجب صياغة أصل السؤال (الأرومة) بلغة سهلة واضحة ومفهومة، حيث لوحظ في كثير من الحالات ان عجز الطالب عن تحديد الاجابة الصحيحة ليس بسبب عدم معرفته للبديل الصحيح، بل لأن أصل السؤال صِيَغ بلغة صعبة أو استخدم عبارات مركبة قد لا يفهمها الطالب.

مثال ضعيف :
- كم عدد السعرات على وجه التقريب التي تَنصح بها لفتاة في الرابعة عشر من عمرها، تزن 45 كيلو غراماً، وذات نشاط عالي في عملها؟
أ- صفر.
ب- 2000
جـ- 2500
د- 3000

صيغة جيدة :

- كم عدد السعرات التي تناسب فتاة في الرابعة عشر من عمرها ووزنها 45 كيلو غرام وذات نشاط معتدل؟

أ- 1500　　　ب- 2000

جـ- 2500　　　د- 3000

ثانياً: القواعد الخاصة بالبدائل:

1- يجب التأكد من وجود بديل صحيح لا خلاف على صحته,أو بديل يمثل أفضل البدائل المطروحة اذا كان الهدف قياس قدرة الطالب على اختيار البديل الأكثر صحةً من بين عدد من البدائل الصحيحة على أن يكون ذلك واضحا" في تعليمات الاختبار.

مثال :

صياغة ضعيفة:

- ما الغرض من الاختبارات الصفية؟
- ثم تكتب مجموعة من البدائل.

صياغة جيدة:

أن الغرض الرئيسي للإختبارات الصفية هو:

- **ثم تكتب مجموعة من البدائل**

2- إجعل جميع البدائل متسقة نحوياً (لغوياً) مع أصل السؤال.

مثال :

صيغة ضعيفة في اللغة الانكليزية:

- The recall of factual information can be measured best with a:
  A- matching item.
  B- multiple- choice item.
  * C- short- answer item.
  D- essay question.

ويلاحظ في المثال السابق عدم إنسجام لغوي بين نهاية أصل السؤال والبديل الرابع (D)، مما يستدعي اما الى اعادة صياغة اصل السؤال او تغيير صيغة البدائل من المفرد الى الجمع.

مثال بصياغة جيدة :

- The recall of factual information can be measured best with:
  A- matching items.
  B- multiple- choice items.
  *C- short- answer items.
  D- essay questions.

مثال بصياغة ضعيفة:

- لإسعاف الشخص المصاب بكسر في الساق يقوم المسعف بـ:

أ- إبقاء المصاب على حاله لحين حضور الطبيب.

ب- خلع ملابس المصاب.

جـ- يُعطي المصاب الشاي والقهوة.

د- وضع جبائر على الساق المكسورة.

ونلاحظ أن البديل الثالث (جـ) لا يتفق لغوياً مع أصل السؤال، لذلك يجب اعادة صياغة هذا البديل ليكون كالآتي:

جـ- اعطاء المصاب الشاي والقهوة.

ومما يجدر ذكره أن الصياغة اللغوية الجيدة لفقرات الاختيارة من متعدد يجب أن تُلاحظ بشكل واضح من خلال قراءة اصل السؤال وكل بديل من البدائل، بحيث يشكلان إنسيابية لغوية واضحة دون توقف اضافة الى الانسجام النحوي.

3- يجب أن تكون جميع البدائل (الصحيحة والخاطئة) متجانسة تعود الى نفس المجال العلمي أو التأريخي والجغرافي وغير ذلك، لأن البديل الذي

لا ينتمي الى ذلك المجال سيكتشفه الطالب ويحذفه من الاختيارات الباقية وبذلك تزداد نسبة تحديد الاجابة الصحيحة بالحرز أو التخمين اي بدون معرفة أو فهمٍ علمي.

مثال:

صياغة ضعيفة:

- أي من العلماء التالية اسمائهم توصل الى صياغة النظرية النسبية؟

أ- نيوتن. ب- آينشتاين.

جـ- أديسون. د- دارون.

ومن هذا المثال يلاحظ أن بامكان الطالب أن يستبعد البديل (جـ) على أساس أن (أديسون) اكتشف الكهرباء ولا علاقة له بالمنظرين الأخرين، وبذلك تزداد نسبة تخمين الاجابة من (25%) الى (33%).

صياغة جيدة:

- أيٍ من العلماء التالية أسمائهم توصل الى صياغة النظرية النسبية؟

أ- نيوتن. ب- اينشتاين.

جـ- دارون. د- فيثاغورس.

مثال آخر:

صياغة ضعيفة:

- اي من الوسائل الاحصائية الوصفية التالية تمثل القيمة الأكثر تكراراً في مجموعة من القيم؟

أ- الوسط الحسابي. ب- الوسيط.

جـ- المنوال. د- تحليل التباين.

وهنا يستبعد الطالب البديل (د) أي تحليل التباين لانه من وسائل الاحصاء الاستدلالي وليس الوصفي.

صياغة جيدة:

- أي من الوسائل الاحصائية الوصفية التالية تمثل القيمة الأكثر تكراراً في مجموعة من القيم؟

أ- الوسط الحسابي.

ب- الوسيط.

جـ- المنوال.

د- الانحراف المعياري.

4- يجب ان تكون بدائل الفقرات جذابة (تبدو للطالب وكأنها إجابة صحيحة حتى يتم اختبار قدرته على اختيار الاجابة الصحيحة المتأكد منها)، ويمكن تصميم البدائل الجذابة من الاخطاء الشائعة في اجابات الطلاب، كأن يستخدم المعلم (فقرات إكمال) لطلبته ويستفيد من اجاباتهم الخاطئة وصياغتها كبدائل جذابة لفقرات الاختيارة من معدد.

5- يجب التأكد من عدم وجود اتفاق لغوي بين أصل السؤال وإحدى البدائل التي تعقبه (لا سيما البديل الصحيح) إذ أن الطالب حتى وإن كان من الطلبة الضعفاء فقد ينتبه إلى ذلك البديل ويختاره حتى وان كانَ غير متأكداً من صحته.

مثال خاطئ:

- أيٍ من الاجهزة التالية يستخدم لقياس درجات الحرارة؟

البارومتر. ب- المانومتر. جـ- المحرار. د- المهزاز.

ويلاحظ أن البديل (د) يتشابه لغوياً مع اصل السؤال، لذلك يجب اعادة صياغة هذه الفقرة على النحو التالي:

- أيٍ من الأجهزة التالية يُستخدم لقياس درجات الحرارة؟

أ- البارومتر. ب- المانومتر. جـ- الترمومتر. د- المهزاز.

6- حاول أن تكون البدائل الصحيحة والخاطئة متساوية في الطول.
لأن طول البديل الصحيح على وجه الخصوص قد يدفع الطالب باختياره اعتماداً على التخمين أكثر من المعرفة او الفهم المسبق.

7 - تجنب استخدام بعض المصطلحات التي كثيراً ما ترتبط مع البدائل الصحيحة من مثل (احياناً sometimes قد may، عادةً usually) لأنها تدفع الطالب الى اختيار ذلك البديل بدون فهم، ويمكن توضيح ذلك في المثال التالي من اللغة الانجليزية:

مثال:

مثال ضعيف:

Lack of attention to learning outcomes during test preparation:

A- will lower the techincal quality of the items.

B- will make the construction of test items more difficult.

C- will result in the grater use of essay questions.

* D- may result in a test that is less relevant to the instructional program.

8 - تجنب استخدام بعض المصطلحات المطلقة مع بعض البدائل وهي (دائماً always، أبداً never، جميع all) لأنها قد تدفع الطالب لاستبعاد البدائل المتضمنة لمثل هذه المصطلحات لمعرفته المسبقة إنها ترتبط بالعبارات الخاطئة.

مثال:

-Achievement tests help students to improve their learning by:

A- encouraging them all to study hard.

*B- informing them of their progress.

C- giving them all a feeling of success.

D- preventing any of them from neglecting their assignments.

ويلاحظ في المثال أعلاه تكرار المصطلح (all) مع البديلين (A,C) اللذان يستبعدهما الطالب مما يزيد من نسبة التخمين الى (50%) ويسهل عليه إختيار البديل الصحيح من البديلين (B,D).

9- عندما تكون البدائل متضمنة مجموعة من الارقام (أعداد، قياسات، تواريخ وغير ذلك) فيفترض ترتيبها ترتيباً منطقياً من الاعلى الى الأدنى وبالعكس.

مثال:

- في أيٍ من الاعوام التالية بدأت الحرب العالمية الاولى:

أ- 1915

ب- 1916

جـ-1918

د- 1919

10- وزع البدائل التي تمثل الاجابات الصحيحة في فقرات الاختبار بشكل عشوائي، ولا تتبع اي نظام قد يكتشفهُ الطالب ويستخدمه للتوصل الى الاجابة الصحيحة بشكل تخمين، أي لا تجعل البديل (جـ)مثلاً هو البديل الصحيح دائماً، بل مرةً (جـ) واخرى (أ)وثالثة (د) وهكذا.

إختبار المطابقة (المزاوجة) Matching Item

وهو أحد الاختبارات التي تتطلب إختيار الإستجابات. ورغم استخدامه بكثرة من قبل معلمي التعليم الاساسي (الابتدائي) لان طريقة إجابته سهلة بالنسبة للأطفال في تلك المرحلة الا إنهُ أقل شيوعاً من إختبارات الصواب والخطأ والاختيار من متعدد. ويتطلب هذا الاختبار من الطالب أن يستخدم معايير العلاقة ليطابق بين كلمات او عبارات تمثل أفكار، مفاهيم، مبادئ، أو أية أشياء أُخرى تعرض على الطالب بشكلٍ قائمتين تسمى القائمة الأولى المقدماتً premises التي تمثل اصل السؤال، أما القائمة الثانية فتُسمى بالاستجابات responses التي تمثل الإختيارات المتوفرة، وكلا القائمتين تؤلفان فقرة الاختبار. إن فقرات هذا الاختبار تصلح لقياس قدرة الطالب على تحديد العلاقات أو التعريفات. كما ان هذا الاختبار يوفر الفرصة للمعلم لتكثيف

محتوىً واسعاً من المادة الدراسية في فقرات محددة لا تأخذ الا وقتاً بسيطاً للاجابة عليها. ومن الميزات الاخرى لهذا النوع انه لا يتطلب وضع بدائل خادعة كتلك التي يتطلبها الاختيار من متعدد، في حين يمثل كل عنصر من قائمة المقدمات (أصل السؤال في الاختيار من متعدد) بينما تكون جميع العناصر في قائمة الاستجابات بدائل متوقعة لذلك العنصر. من ناحية اخرى يمكن القول ان فقرات هذا الاختبار يمكن اشتقاقها من فقرات الاختيار من متعدد عندما يتضح تكرار نفس البدائل في عدد من فقرات الاختيار من متعدد (Gronlund,2006)، والأمثلة الآتية توضح ذلكً:

مثال:

1- أي من الفقرات التالية هي الأقل استخداماً في التشخيص التربوي؟
أ- الاختيار من متعدد. ب- الصواب والخطأ.
جـ- الاسئلة القصيرة. د- المزاوجة.

2- أي من الفقرات التالية تقيس مخرجات تعليمية متنوعة؟
أ- الاختيار من متعدد. ب- الصواب والخطأ.
جـ- الاسئلة القصيرة. د- المزاوجة.

3- أي من الفقرات التالية يصعب تصحيحها بشكل موضوعي؟
أ- الاختيار من متعدد. ب- الصواب والخطأ.
جـ- الاسئلة القصيرة. د- المزاوجة.

4- اي من الفقرات التالية توفر فرصة أكبر للتخمين؟
ا- الاختيار من متعدد. ●ب- الصواب والخطأ.
جـ- الاسئلة القصيرة. د- المزاوجة.

5- أي من الفقرات التالية تُقيس قدرة الطالب على تحديد العلاقة بين المتغيرات؟
أ- الاختيار من متعدد. ب- الصواب والخطأ.
جـ- الاسئلة القصيرة. ●د- المزاوجة.

ومن الأمثلة أعلاه يمكن صياغة فقرة لاختبار المطابقة بتجميع أصل السؤال في جميع الفقرات لتمثل قائمة (المقدمات),ووضع البدائل الاربعة لتمثل قائمة(الاستجابات) وكما يأتي:

المقدمات	الاستجابات
ب . . 1- الأقل استخداماً في التشخيص التربوي.	أ- الاختيار من متعدد.
أ . . 2-يُقيس مخرجات تعليمية متنوعة.	ب- الصواب والخطأ.
جـ . . 3- يصعب تصحيح اجاباته بشكل موضوعي.	جـ- الاسئلة القصيرة.
ب . . 4- يُوفر فرصة أكبر للتخمين.	د-المزاوجة.
د . . 5- يُقيس العلاقة بين المتغيرات.	

ميزات اختبار المطابقة:

1- يمكن اعداده بسهولة لذلك يرغب المعلمون في استخدامه.

2- أن مجال الحزر والتخمين في هذا النوع أقل مما في اختبار الصواب والخطأ والاختيار من متعدد ويقل الحزر كلما زاد عدد الاستجابات.

3- يمكن تقدير علاماته بموضوعية كاملة كما في الانواع الموضوعية الاخرى.

4- يشجع الطالب على اكتشاف وتحديد العلاقات بين المتغيرات في جميع المواد الدراسية.

عيوبه:

على الرغم من الميزات الايجابية التي ذكرت في اعلاه فان هناك بعض العيوب التي يتصف بها هذا النوع ومنها:

1- لكي يكون فعالاً يتطلب من المعلم ايجاد أمثلة عديدة للعلاقات المتناضرة مما يصعب توفيرها من قبل المعلم في بعض الاحيان فيلجأ الى تصميم أنواع اخرى لا سيما الإكمال والصواب والخطأ، او وضع عناصر في كلا القائمتين لا إنسجاماً علمياً بينها فيضعف من كفاءته.

2- يصعب تصميم فقرات تقيس مستويات عليا من التفكير مثل التحليل، التركيب والتقويم مما يحصر استخدامه على المستويات الدنيا.

3- يحتل مكاناً أكبر في ورقة الاختبار، ويحتاج الى جهد طباعي أكبر من قبل المعلم. وبالنسبة للطالب يحتاج الى جهد ووقت أكبر في قراءة فقراته.

قواعد تصميم فقرات المطابقة:

1- يُفترض بالمعلم عند عرضه لفقرات هذا الاختبار في ورقة الاسئلة الامتحانية أن يضع في الجهة اليمنى قائمة المقدمات التي تسبقها خطوط منقطة ويرقم تسلسل مفردات هذه القائمة، وأن يرتب قائمة الاستجابات في الجهة اليسرى من الورقة وأن تُسبق عناصرها بحروف الالفباء. ويرشد الطالب بأن يضع الحرف الدال على الاختيار الصحيح من قائمة الاستجابات على الخط المنقط الموجود أمام الرقم الخاص الذي يتطابق مع هذا الاختيار في قائمة المقدمات وكما موضح في أدناه:

مثال:

طابق بين كل نهر من الانهار الموجودة في القائمة (ب) مع البلد الذي يجري فيه من القائمة (أ) وضع الحرف الدال عليه على الخط الموجود أمام رقم ذلك البلد.

المقدمات	الاستجابات
هـ (1) العراق.	أ- الليطاني.
د (2) مصر.	ب- نهر اليرموك.
جـ (3) سوريا.	جـ- بَرَدى.
أ (4) لبنان.	د- النيل.
	هـ- الفرات.

2- تجنب المطابقة التامة، اي التي تحدث نتيجة تساوي عدد العناصر في القائمة (أ) مع عدد العناصر في القائمة (ب)، لان ذلك يدفع بعض الطلبة لإستخدام التخمين في إجابة الفقرة التي لا يعرفون مطابقتها حيث يتم تركها الى ما بعد إجراء المطابقات الأخرى، وبالتالي يسمى عنصراً في القائمة (ب) تتم مطابقته مع العنصر المتبقي في القائمة (أ) بشكل آلي وتكون المطابقة في الغالب صحيحة.

3- يجب أن تطبع عناصر قائمتي المقدمات والاستجابات في نفس الصفحة وعلى جانبيها بحيث لا نجعل الطالب مضطراً لان يقرأ قائمة المقدمات في صفحة، ثم يقلب الصفحة ليقرأ قائمة الاستجابات لان في ذلك تشويش على الطالب واستهلاك للوقت، ومتعب للطلبة الصغار على وجه الخصوص.

4- يمكن صياغة فقرات المطابقة لتقيس قدرة الطالب على تصنيف الاشياء، مثل تصنيف النباتات، وتصنيف المواد الغذائية، وتصنيف الحيوانات، واعضاء جسم الانسان وغير ذلك وكما نلاحظ في الامثلة الآتية:

مثال (1):

أ (١). كربوهيدرات.	أ- أرز.
د (٢). دهنيات.	ب- بيض.
ب (٣). زلاليات.	جـ- لحم الدجاج.
جـ (٤). بروتينات.	د- زبدة.
	هـ- تفاح.

مثال (2):

القائمة (أ) أجهزة الجسم	القائمة (ب) الاعضاء
(أ) ...... (1) جهاز التنفس	أ- الحويصلات الهوائية.
(ج) ...... (2) جهاز البولي	ب- الشريان الأبهر.
(د) ...... (3) جهاز الهظم	جـ- الحالب.
(ب) ..... (4) جهاز الدوران	د- القولون.
	هـ- الترقوة.

5- من الضروري أن تكون عناصر كل قائمة متجانسة، لأن العناصر غير المتجانسة مع العناصر الأخرى للقائمة قد يستغلها الطالب للوصول الى الاجابة الصحيحة بدون فهم أو معرفة. فلو أردنا اختبار قدرة الطالب على معرفة العلاقة بين مخترعات وعلماء قاموا باختراعها، وأوردنا عالماً لم يكن مخترعاً بل طبيباً أو منظراً، فإن الطالب قد يكتشف ذلك بسهولة ويصل الى الاجابة من خلال الاختلاف بين ذلك العالم والمخترعين الآخرين في القائمة.

مثال (3):

أ (1). بل.	أ- جهاز الهاتف.
ب (2). ماركوني.	ب- الراديو.
د (3). أديسون.	جـ- التحليل النفسي.
جـ (4). فرويد.	د- الكهرباء.
	هـ- الحاسوب.

ونلاحظ أن فرويد لم يكن مخترعاً، ومصطلح التحليل النفسي ليس من المخترعات الصناعية، مما يدفع الطالب بالمطابقة بين هذين العنصرين رغم عدم معرفته بعالم النفس (فرويد).

6- لا تجعل عناصر قائمة المقدمات طويلة من حيث عدد عناصرها بحيث تتعب الطالب عند قراءتها ويمكن تحديد عدد عناصرها بخمسة أو أقل من عشرة، وإختصر في كتابة الاستجابات بحيث لا يكون كل عنصر منها طويلاً لأن الطالب يحتاج الى قراءتها عدة مرات ليختار منها العنصر الذي يتطابق مع عنصراً مناسباً له في قائمة المقدمات.

7- حاول أن ترتب عناصر كلا القائمتين (وخاصة قائمة الاستجابات) ترتيباً ابجدياً أو منطقياً عندما تكون لديك أرقام تمثل تواريخ أو احصاءات معينة، لان ذلك يساعد الطالب على اختيار الاجابة الصحيحة بشكل أسهل.

8- على الرغم مما ذُكر بتركيز استخدام هذا النوع على قياس قدرة الطالب في مجال المعرفة أو التصنيف، فان بالامكان استخدامه لقياس القدرة على فهم وتطبيق بعض المبادئ أو القواعد اللغوية، فإذا أراد المعلم أن يقوّم قدرة طلابه على التمييز في تطبيق استخدام الأسماء والأفعال والصفات والظروف وفي الخرائط الجغرافية، مثلا فان بامكانه تحقيق ذلك من خلال فقرات هذا الاختبار وكما في المثال الآتي: مثال في اللغة العربية.

مثال في اللغة العربية:

أ (1). فعل.	أ- ساعد عليٌّ أخاهُ.
جـ (2). صفة.	ب- نسيتُ الكتابَ فوقَ المنضدة.
ب (3). ظرف.	جـ- محمدٌ تلميذٌ مجتهدٌ.
د (4). إسم	د- جاءَ أحمدُ ماشياً.
	هـ- ذهبَ الطالبُ إلى المدرسةِ.

مثال في الجغرافية:

المقدمات (أسماء أنهار)	الاستجابات (ترقيم مواقع الانهار)
(هـ) (1). النيل.	
(أ) (2). دجلة.	
(ج) (3). الليطاني.	
(ب) (4). الفرات.	

إدارة وتصحيح الأختبار Administering and Scoring the Test

بعد أن ينتهي المعلم من تصميم الاختبار بشكله النهائي يبدأ بتطبيقه على المجموعات التي يهدف الى قياس تحصيلها بواسطة ذلك الاختبار. ولأجل تحقيق الاهداف التي وضع من أجلها ذلك الاختبار علينا أن نهتم باسلوب ادارته وتطبيقه، آخذين بنظر الاعتبار ضبط جميع الظروف المادية والنفسية التي تشكل الجو الامتحاني مثل خفض الضوضاء واساليب التشويش أثناء اجابة الطالب، وضرورة أن تكون تعليمات الاختبار واضحة ومفهومة، وترك مسافات كافية بين مقاعد جلوس الطلبة لمنع عملية الغش. وينصح بعض الباحثين ضرورة توضيح تعليمات الاختبار بشكل شفوي رغم انها مكتوبة في ورقة الامتحان ويمكن استخدام السبورة لتوضيح بعض أساليب الاجابة على الاسئلة لتكون هذه الاجراءات واضحة أمام الطالب ولا تدفعه للانشغال بها دون تركيز اهتمامه للاجابة عن اسئلة الاختبارات.

ومن أجل تسهيل عملية تصحيح أوراق الإجابة، على المعلم أو المصحح ان يجعل مكان الاجابات في ورقة الامتحان على جهة واحدة (اليمنى واليسرى) مثلا لتسهيل عملية التصحيح. وكما موضح في المثال التالي:

مثال:

فقرات الصواب والخطأ:

(ص)(خ). يتكون الماء من اتحاد ذرّة من الهيدروجين وذرتان من الأوكسجين.

فقرات اختيار من متعدد:

جـ أي من المدن الأتية هي عاصمة دولة قطر؟

(أ) الكويت. (ب) بغداد. (جـ) الدوحة. (د) مسقط.

وفي حالة وجود ورقة منفصلة للاجابة فإنها مطبوعة في العادة بشكل يحقق هذه القاعدة، وذلك ليسهل تصحيحها يدوياً. وهناك صورة متعددة لتصحيح الإجابات يدوياً مثل المفتاح المثقب والمفتاح الشفاف ومفتاح الكربون. وبعد إنتشار الحاسوب في المجال التربوي ظهرت أنواع معيارية لإستمارات الإجابة Standard answer sheet باستخدام الكمبيوتر حيث يطلب من الطالب ملء دوائر خاصة فارغة على ورقة الاجابة بقلم رصاص محدد لهذا الغرض وتقوم الحاسبة بالمسح الضوئي Optical Saning للاجابات الصحيحة وتعطي النتائج النهائية للطلبة بوقت قصير ودون بذل المعلم للجهد الشاق الذي يتطلبه التصحيح اليدوي التقليدي.

ومن الامور المادية (الفيزيائية) المطلوب الاهتمام بها اثناء تطبيق الاختبار إضاءة الصف الكافية، ودرجة حرارة الفصل الذي يجري فيه الاختبار، وكذلك تحديد الوقت الخاص بالاختبار لكي يكون الطالب على علم بها ولا باس باعلانها للطلبة أو كتابتها بشكل واضح في اعلى الصفحة الأولى من الورقة الامتحانية. وأخيراً لابد من الاعلان للطلبة قبل بدء اجاباتهم ما اذا كان المعلم سيستخدم معادلة ازالة أثر التخمين.

تحليل فاعلية فقرات الاختيار

بعد تطبيق الاختبار على المجموعة المستهدفة قياس تحصيلها أو جزء منها، يتم تحليل فقرات الاختبار بناء على استجابات أفراد المجموعة على تلك

الفقرات من أجل تحديد كفاءة او فاعلية تلك الفقرات. ويمكن التأكد من تلك الفاعلية بعدة طرق منها الطريقة العادية حيث تُفتح مناقشات شفوية مع أفراد تلك المجموعة عندما يكون عددهم غير كبير (كأن تكون مجموعة صفية) وذلك باستعراض الفقرات واحدةً واحدة والاستماع الى رأي المجموعة عن كل فقرة من حيث محتوى الفقرة وصياغتها ووضوحها أو غموضها وكذلك من حيث كفاءتها في قياس الهدف السلوكي الذي وضعت لقياسه أو من حيث مستوى صعوبتها وسهولتها، ويمكن لمعلم الصف أن يرد على طلبته في الموافقة او تبيان وجهة نظر مخالفة لوجهة نظر المجموعة. ان هذا الاسلوب وان كان اسلوباً نوعياً إلا أن له الكثير من الفوائد التي تعود على المعلم عندما يعيد النظر في الاختبار نفسه او عندما يفكر في تطوير اختباراً آخر في المستقبل، بالاضافة الى المردود الايجابي الذي يكسبه الطلبة من ذلك التطوير.

أما الاسلوب الآخر لتحليل استجابات المجموعة على فقرات الاختبار فهو الاسلوب الكمي الذي يسمى بتحليل الفقرات item analysis فيشمل مؤشرات عدة منها مستوى صعوبة الفقرة item difficulty ومعامل تمييز الفقرة item discrimination اللذان سنتحدث عنهما في الصفحات القادمة إضافة الى مؤشرات اخرى من مثل فاعلية البدائل الخاطئة.

إن تحليل استجابات الممتحنين على الاختبار (ولا سيما الاختبار الموضوعي) يؤدي بالاضافة الى ما ذُكرَ اعلاه غرضين مهمين هما:

1- ان معرفة استجابات الطلبة لفقرات الاختبار يفيد في توفير معلومات تشخيصية لبيان مدى التعلم الحاصل في التدريس الصفي من عدمه.

2- توجيه إسلوب التدريس في المستقبل بما يتلائم مع قدرات الطلبة وزيادة استعداداتهم لمواجهة الاختبارات وكيفية الاجابة على فقراتها.

ومن أجل تحديد مستويات الصعوبة والتمييز والحكم على البدائل التي تتضمنها فقرات الاختبار نحتاج الى معلومات عن تحديد الفئتين العليا (وهي الفئة التي تمثل مجموعة الطلبة الذين حصلوا على أعلى العلامات الكلية في الاختبار) أي ابتداءً من أعلى علامة في المجموعة نزولاً، والفئة الدنيا (التي

تمثل مجموعة الطلبة الذين حصلو أقل العلامات ابتداءً من أقل علامة كلية تَمَّ تحقيقها في المجموعة صعوداً). وعادة ما يؤخذ من كل مجموعة (25%، أو 27%، أو 33%). كما نَحتاج لاجراء هاتين العمليتين معرفة توزيع استجابات الطلبة على كل فقرة وفقاً للمثال الآتي:

مثال:

- أي من الدول العربية التالية تقع في قارة افريقيا؟

أ- سلطنة عُمان.

ب- العراق.

●جـ- مصر.

د- لبنان.

ادناه نتائج استجابات الطلاب من الفئتين العليا والدنيا على الفقرة أعلاه:

البدائل	أعداد الطلبة حسب اختياراتهم	
	المجموعة العليا 25%	المجموعة العليا 25%
أ-	1	2
ب-	1	3
جـ- (البديل الصحيح)	10	6
د-	3	5
المجموع	15	15

صعوبة الفقرة: Item difficulty

ويقصد بها مستوى التعقيد الذي يواجههُ الطالب في الاجابة الصحيحة على الفقرة الاختبارية، وما إذا كان ذلك المستوى عالياً أم متوسطاً. ومن الطبيعي ان سهولة الفقرة هي عكس صعوبتها اي ان الفقرة السهلة هي تلك التي بامكان اغلبية الطلبة تحديد الاجابة الصحيحة عنها دون مشكلة لانها لا تمثل مستوى من التعقيد الذي تتميز به الفقرة الصعبة. ويحدد مستوى صعوبة الفقرة اجرائياً بالنسبة المئوية للطلبة الذَين حققوا الاجابة الصحيحة على تلك الفقرة ويمكن تمثيله بالمعادلة التالية:

$$\text{م ص} = \frac{\text{عدد الطلبة الذين أجابوا على الفقرة إجابة صحيحة}}{\text{العدد الكلي للطلبة الممتحنين}}$$

مثال:

لو كان مجموع الطلاب الذين أجابوا على الفقرة (25) طالباً فقط من مجموع (50) ممتحناً، فان معامل صعوبة هذه الفقرة = $\frac{25}{50} = 0.50$

وهناك من يعتمد في ايجاد معامل صعوبة الفقرة على تحديد الطلاب الذين أجابوا إجابة صحيحة من الفئة العليا للمجموعة، مضافاً اليها عدد الذين اجابوا اجابة صحيحة على الفقرة من أفراد المجموعة الدنيا، وبذلك فان معامل الصعوبة وفقاً لهذه الطريقة يتطلب من المصحح أو المعلم القيام بالاجراءات الآتية:

أ- تصحيح اجابات المجموعة التي أدت الامتحان.

ب- ترتيب اوراق الاجابات بشكل تنازلي وفقاً للدرجة الكلية التي حصل عليها كل طالب من المجموعة أي من اعلى درجة الى أدناها.

جـ- تحديد المجموعة العليا من الطلاب الذين حصلوا على أعلى الدرجات في الاختبار، وتحديد المجموعة الدنيا من الطلاب الذين حصلوا على أدنى الدرجات. وهناك من يقترح تحديد نسبة كل من المجموعتين بمقدار 25%، أو 27%، أو 33% من العدد الكلي لأفراد المجموعة ونستبعد الفئة التي تقع بين المجموعتين العليا والدنيا (أي الفئة الوسطى) خارج العملية.

د- تحديد عدد الطلاب الذين أجابوا على الفقرة التي نريد تحديد مستوى صعوبتها في المجموعتين العليا والدنيا معاً لتقسم الناتج على العدد الكلي لافراد المجموعتين مضروباً في مائة.

مثال:

لنفرض أن (40) طالباً طبق عليهم إختباراً تحصيلياً في مادة دراسية معينة. وبعد تصحيح اجاباتهم تَمَّ تحديد المجموعتين العليا والدنيا بنسبة (25%) لكل منهما اي (10) طلاب لكل مجموعة. ثُمَّ تحديد عدد الطلاب الذين اجابوا

إجابة صحيحة على الفقرة (1) من الاختبار (على سبيل المثال) فوجد أن عددهم في المجموعة العليا كان (8) طلاب، وفي المجموعة الدنيا أجاب على نفس الفقرة إجابة صحيحة (2) طالب. فما هو معامل صعوبة تلك الفقرة؟

الحل:

$$\text{م ص} = \frac{2 + 8}{10 + 10}$$

$$\text{م ص} = \frac{10}{20}$$

م ص= 0.50 معامل صعوبة الفقرة رقم (1) من الاختبار.

ويضاف إلى الأسلوبين السابقين اسلوب ثالث لايجاد معامل صعوبة الفقرة يُحسب كالآتي :

$$\text{معامل الصعوبة} = \frac{\text{مجموع الدرجات التي حصل عليها أفراد المجموعة على الفقرة}}{\text{عدد الطلاب الكلي} \times \text{الدرجة المخصصة للسؤال أو الفقرة}}$$

ومن الممكن استخدام هذا الاسلوب في حالة تحديد صعوبة السؤال المقالي أيضاً. (عيد، 2007)

أما معيار الحكم على مدى صعوبة الفقرة أو سهولتها فقد إتفق أغلب الباحثين على تحديده بـ (0.50)، ولكن لا بأس أن يمتد بين (0.40- 0.60) حيث يُشير بعض الباحثين ان هذا المدى لمستوى صعوبة الفقرة يزيد من مستوى تمييزها (علام، 2006) وعليه فإن النسبة إذا إزدادت لتقترب من (1) تكون سهلة وكلما قلت لتقترب من (صفر) فانها تكون صعبة، ومن ذلك يمكن الاستنتاج أن الفقرة التي معامل صعوبتها (0.90) هي فقرة سهلة جداً، وتلك التي معامل صعوبتها (0.15) هي فقرة صعبة جداً.

وبالاضافة للأساليب السابقة الخاصة بايجاد معامل صعوبة الفقرة هناك من يأخذ بعض العوامل الاخرى التي لها علاقة باجابة الطلبة على الفقرة بنظر الاعتبار في حساب معامل الصعوبة، لا سيما عامل التخمين gussing وعامل

ترك الاجابة عن الفقرة من قبل بعض الطلبة، فلإيجاد معامل صعوبة الفقرة باستخدام معادلة إزالة اثر التخمين يمكن ملاحظة المثال الأتي:

مثال:

لو فرضنا إن (70) من الطلبة المتحنين أجابو على إحدى الفقرات في إختبار معين إجابة صحيحة، بينما أجاب (30) منهم إجابة خاطئة على نفس الفقرة ذات البدائل الأربعة. فما هو معامل الصعوبة باستخدام معادلة التخمين؟.

معامل الصعوبة باستخدام معادلة التخمين = $\frac{\text{ص} - \frac{\text{خ}}{\text{ن} - 1}}{\text{ص} + \text{خ}}$

حيث أن: ص= عدد الطلبة الذين أجابوا على الفقرة إجابة صحيحة.

خ= عدد الطلبة الذين أجابوا على الفقرة إجابة خاطئة.

ن= عدد بدائل الإجابة.

معامل الصعوبة = $\frac{70 - \frac{30}{4 - 1}}{70 + 30}$

= $\frac{70 - 10}{100}$ = 0.60 معامل صعوبة الفقرة.

أما بالنسبة لأخذ عدد الطلبة الذين تركوا الفقرة دونَ أن يجيبوا عليها بنظر الاعتبار فيمكن توضيحهُ بالمثال الآتي:

مثال:

لنفرض ان عدد أفراد المجموعة التي أجابت على الاختبار كان (300) طالب أجاب منهم إجابة صحيحة على فقرة ما (150) طالباً وأخطأ فيها (120) طالب بينما ترك الاجابة عن الفقرة (30) طالباً، فما هو معامل صعوبة تلك الفقرة إذا علمتَ أن عدد بدائل الفقرة هو (5)؟

الحل:

في هذه الحالة نطبق المعادلة الآتية:

$$\text{مستوى الصعوبة} = \frac{\text{ص} - \frac{\text{خ}}{\text{ن} - 1}}{\text{ح} - \text{ك}}$$

حيث أن: ص= عدد الطلاب الذين أجابوا اجابة صحيحة على الفقرة.
خ= عدد الطلاب الذين أجابوا اجابة خاطئة على الفقرة.
ن= عدد بدائل الاجابة على الفقرة.
ح = العدد الكلي للطلبة.
ك= عدد الطلبة الذين تركوا الاجابة على الفقرة.

$$\text{معامل صعوبة الفقرة} = \frac{150 - \frac{120}{5 - 1}}{300 - 30} = 0.40$$ معامل الصعوبة.

وأخيراً لا بد أن يتساءل البعض لا سيما العاملين في الميدان التربوي هل ان مستوى الصعوبة المقبول بحدود (0.50) يجب ان نحافظ على تحقيقه في جميع فقرات الاختبار؟ اليس من حق المعلم أن يضع مجموعة فقرات صعبة يحدد من خلالها الطلاب ذوي القدرات العليا؟ وللأجابة على هذا السؤال لا بد من القول أن الاختبار التحصيلي لا بد أن يضم تدرجاً واسعاً من مستويات صعوبة فقراته. ولأجل أن يكون واضع الاختبار أو معلم الصف متوازناً في ذلك ينصح بعض الباحثين أن يضم الاختبار حوالي:
50% من الأسئلة ذات مستويات صعوبة تتراوح بين 0.25 - 0.75
25% من الاسئلة ذات مستويات صعوبة اعلى من 0.75
25% من الاسئلة ذات مستويات صعوبة اقل من 0.25. (Kubisyn & Borch, 1987)

تمييز الفقرات Item discrimination

لأجل أن نوضح معنى تمييز الفقرة لا بد ان تستحضر مثالاً بسيطاً من استخدام الأجهزة العلمية وأهميتها في القياس الفيزيائي والقياس الجسمي أيضاً ولماذا نثق في نتائج تلك الأجهزة. مما لا شك فيه اننا نثق في المحرار - مثلاً - اذا ما قدم لنا قراءات صحيحة عندما نقيس به درجة حرارة الغرفة أو

درجة حرارة مريض معين. ولا شك أن درجات الحرارة تتفاوت من مكان الى آخر ومن مريضٍ لمريض، فان لم تتباين القراءات على هذا المحرار فاننا بكل تأكيد لا نثق بتلك القراءات لانها لا تميز مثلاً بين المكان الحار والمكان الأقل حرارةً. وكذلك الحال بالنسبة الى جهاز قياس ضغط الدم لدى الانسان لا يمكن ان نثق به اذا ما أعطانا نفس مستوى ضغط الدم لكل من هم في مرحلة الشباب، واولئك الافراد الذين تجاوزوا سن الخمسين من أعمارهم، لاننا نعرف أساساً ان مستوى ضغط دم الإنسان يزداد كلما تقدم الانسان في العمر. إن عدم الثقة هذه ناشئة من أن هذا الجهاز (أو المقياس) لا يميز بين منخفضي الضغط ومرتفعي الضغط، ونفس هذه الحالة تنطبق على الاختبارات بشكل عام والاختبارات التحصيلية بشكل خاص، فالاختبار الجيد هو الذي يميز بين الاداء العالي والاداء المنخفض، أن مستوى تمييز الفقرة item dis-crimination يعني الدرجة التي تحدد قدرة الفقرة الاختبارية على التمييز بين أداء الطلبة من ذوي المستوى العالي وأداء الطلبة من ذوي المستوى المنخفض، وان هذه الخاصية تُعتبر واحدة من المواصفات الهامة المفروض توفرها في فقرات الاختبار الجيد. اما من الناحية الاجرائية فإن معامل تمييز الفقرة يمكن إيجاده وفق الخطوات الآتية:

1 - اتباع نفس الخطوات الاولى المذكورة لايجاد معامل صعوبة الفقرة من حيث ترتيب أوراق إجابات جميع الطلبة المتحنين من أعلى درجة كلية الى اقل درجة كلية، واخذ (27%) أو (25%) من قمة وقاعدة العلامات لتحديد عدد أفراد المجموعة العليا والدنيا، وكم منهما أجاب على الفقرة اجابة صحيحة.

2- تطبيق معادلة إيجاد معامل تمييز الفقرة التالية:

معامل تمييز الفقرة (م ت)=

$$\frac{\left(\text{عدد طلاب المجموعة العليا ذوي الإجابات الصحيحة على الفقرة}\right) - \left(\text{عدد طلاب المجموعة الدنيا ذوي الإجابات الصحيحة على الفقرة}\right)}{\text{نصف عدد أفراد المجموعتين}}$$

ومن المثال السابق الذكر الخاص بإيجاد معامل صعوبة الفقرة يكون معامل تمييز تلك الفقرة:

$$\text{م ت} = \frac{8 - 2}{10}$$

م ت = 0.6 معامل التمييز

أما المعيار الممكن استخدامه للحكم على معامل التمييز المقبول للفقرة، فيمكن الاعتماد على رأي (Ebel, 1979) الذي اقترح القواعد التالية للحكم على الفقرة:

1- تحذف الفقرة التي يكون معامل تمييزها سالب.
2- الفقرة ذات معامل التمييز الذي يتراوح بين (صفر- 0.19) تعتبر ضعيفة وتحذف.
3- الفقرة التي معامل تمييزها بين (0.20- 0.39) تعتبر فقرة مقبولة.
4- الفقرة التي معامل تمييزها اعلى من (0.39) تُعتبر فقرة جيدة.

وبالاضافة الى الاسلوب الموضح سلفاً المستخدم لايجاد معامل تمييز الفقرة فهناك اسلوب آخر لايجاده من خلال استخدام اختبار "ت" t.test للتأكد من دلالة الفروق بين متوسطَي علامات طلاب المجموعتين العليا والدنيا، فاذا وجدت فروق ذات دلالة إحصائية بمستوى لا يزيد عن (0.05) فان ذلك يعني أن الفقرة مميزة، والعكس صحيح.

مثال:

طبق اختبار تحصيلي يتضمن (20) فقرة موضوعية على مجموعة من الطلاب عددهم (40) طالباً، وبعد ترتيب علاماتهم الكلية على الاختبار ثمَّ أخذ (25%) من أعلى العلامات لتمثل علامات المجموعة العليا، و (25%) من أدنى العلامات لتمثل علامات المجموعة الدنيا. أوجد الفروق بين متوسطي علامات المجموعتين على الفقرة الأولى من الاختبار لتحديد ما اذا كانت تلك الفروق دالة إحصائياً بمستوى 0.05 للحكم على تميز تلك الفقرة.

علامات المجموعة العليا على الفقرة الاولى	علامات المجموعة الدنيا على الفقرة الاولى
1	صفر
1	1
1	صفر
صفر	صفر
1	1
1	1
صفر	صفر
1	1
1	صفر
1	1

وعند مقارنة القيمة التائية المحسوبة (20.8) مع القيمة التائية النظرية وبدرجة حرية مقدارها (10+10-2) أي (18) عند مستوى (0.05) والتي تبلغ (2.101) نجد ان القَيمة المحسوبة اكبر من القيمة النظرية، أي أن:

20.8 > 2.101 وعليه نرفض الفرضية الصفرية التي مفادها:

لا توجد فروق ذات دلالة إحصائية بمستوى (0.05) بين متوسط علامات المجموعة العليا ومتوسط علامات المجموعة الدنيا، أي أن الفقرة (1) مميزة بين المجموعتين العليا والدنيا.

ونعمل الشيء نفسه مع البقية فقرات الاختبار لنقرر أي فقرة تُحذف واي فقرة تبقى. علماً بأن الفقرة المميزة هي التي تكون فيها قيمة (ت) المحسوبة اعلى من القيمة النظرية لتكون الفروق دالة بين متوسطي علامات المجموعتين.

وبالاضافة لما تم توضيحه من أساليب ايجاد معامل تمييز الفقرة، هناك إسلوب آخر يَتميز بالبساطة ويعطي نتائج مقاربة لتلك الناشئة عن استخدام الاساليب الاخرى من خلالَ إستخدام جداول فلانجن Flangin Table حسب الخطوات التالية:

1- نحسب نسبة الأفراد الذين أجابوا على الفقرة إجابة صحيحة من الفئة العليا، ونسبة الافراد الذين أجابوا على نفس الفقرة إجابة صحيحة من الفئة الدنيا (بعدَ ترتيب أوراق اجابات الطلبة وفقاً لدرجاتهم الكلية على عموم الاَختبار وتحديد 27% أو 25% من الاوراق لتمثل الفئة العليا والفئة الدنيا).

2- من جدول فلانجن الخاص بتحديد معامل تمييز الفقرة نحدد موقع الفئة العليا على المحور الافقي العلوي للجدول، ونحدد موقع الفئة الدنيا على المحور العمودي الجانبي من الجهة اليمنى. ومن خلال مد خطين عمودي وأُفقي من الاعلى والجانب انطلاقاً من نسبتي عدد طلاب الفئتين العليا والدنيا الذين أجابوا على الفقرة إجابة صحيحة يتم تحديد معامل التمييز الناشئ عن التقاء الخطين، وكما موضح في المثال الآتي:

مثال:

جد معامل تمييز الفقرة السابعة من اختبار مُفترض اجاب عليها بشكل صحيح (14) طالباً من أصل (20) يمثلون طلاب الفّئة العليا، بينما أجاب عليها بشكل صحيح (10) طلاب من أصل (20) من طلبّة المجموعة الدنيا، باستخدام جداول فلانجن.

الحل:

1- نسبة طلبة الفئة العليا الذين اجابوا على الفقرة بشكل صحيح=
= 0.70

2- نسبة طلبة الفئة الدنيا الذين أجابوا على الفقرة بشكل صحيح=
= 0.50

3- من جدول فلا نجن الموجودة في الصفحة التالية نمد خطاً من الرقم (70) على الخط الافقي العلوي للجدول وخطاً اخر من الرقم (50) على الخط العمودي الأيسر للجدول، نلاحظ أن العمودين يلتقيان عند الرقم (21) من الارقام الموجودة داخل الجدول حيث يمثل هذا الرقم معامل تمييز الفقرة.
ولو قمنا باستخدام المعادلة السابقة وهي:

$$\text{م ت} = \frac{14 - 10}{20}$$

م ت = 0.20 قيمة معامل تمييز الفقرة

ونلاحظ ان معامل التمييز باستخدام جدول فلانجن والذي مقداره (0.21)، لا يختلف كثيراً عن هذا المعامل الذي مقداره (0.20) ومن خلال مقارنة القيمتين مع معيار قبول التمييز الذي حدده (إيبل) والموضح في الصفحات السابقة نرى ان المعاملين مقبولين.

جدول فلانجن لحساب معامل تمييز الفقرات

الفئة الأدنى

	٢	٦	١٠	١٤	١٨	٢٢	٢٦	٣٠	٣٤	٣٨	٤٢	٤٦	٥٠	٥٤	٥٨	٦٢	٦٦	٧٠	٧٤	٧٨	٨٢	٨٦	٩٠	٩٤	٩٨
٢	٠٠	١٩	٣٠	٣٧	٤٣	٤٨	٥١	٥٥	٥٨	٦١	٦٣	٦٦	٦٨	٧٠	٧٢	٧٣	٧٥	٧٧	٧٩	٨٠	٨٢	٨٤	٨٦	٨٨	٩١
٦		٠٠	١١	١٩	٢٦	٣١	٣٦	٤٠	٤٤	٤٧	٥٠	٥٣	٥٦	٥٩	٦١	٦٤	٦٦	٦٨	٧١	٧٣	٧٦	٧٨	٨١	٨٤	٨٨
١٠			٠٠	٠٨	١٥	٢١	٢٦	٣٠	٣٤	٣٨	٤١	٤٥	٤٨	٥١	٥٤	٥٧	٦٠	٦٣	٦٥	٦٨	٧١	٧٤	٧٧	٨١	٨٦
١٤				٠٠	٠٧	١٢	١٨	٢٢	٢٧	٣١	٣٤	٣٨	٤٢	٤٥	٤٨	٥١	٥٤	٥٧	٦٠	٦٣	٦٧	٧٠	٧٤	٧٨	٨٤
١٨					٠٠	٠٦	١١	١٦	٢٠	٢٥	٢٨	٣٢	٣٦	٣٩	٤٣	٤٧	٤٩	٥٣	٥٦	٦٠	٦٣	٦٧	٧١	٧٦	٨٢
٢٢						٠٠	٠٦	١٠	١٥	١٩	٢٣	٢٧	٣١	٣٤	٣٨	٤٢	٤٥	٤٩	٥٢	٥٦	٦٠	٦٣	٦٨	٧٣	٨٠
٢٦							٠٠	٠٥	٠٩	١٤	١٨	٢٢	٢٦	٣٠	٣٣	٣٧	٤١	٤٤	٤٨	٥٤	٥٦	٦٠	٦٥	٧١	٧٩
٣٠								٠٠	٠٤	٠٩	١٣	١٧	٢١	٢٥	٢٩	٣٣	٣٧	٤٠	٤٤	٤٩	٥٣	٥٧	٦٣	٦٨	٧٧
٣٤									٠٠	٠٤	٠٩	١٣	١٧	٢١	٢٥	٢٩	٣٣	٣٧	٤١	٤٥	٤٩	٥٤	٦٠	٦٦	٧٥
٣٨										٠٠	٠٤	٠٨	١٣	١٦	٢٠	٢٥	٢٩	٣٣	٣٧	٤٢	٤٧	٥١	٥٧	٦٤	٧٣
٤٢											٠٠	٠٤	٠٨	١٢	١٦	٢٠	٢٥	٢٩	٣٣	٣٨	٤٣	٤٨	٥٤	٦١	٧٢
٤٦												٠٠	٠٤	٠٨	١٢	١٦	٢١	٢٥	٣٠	٣٤	٣٩	٤٥	٥١	٥٩	٧٠
٥٠													٠٠	٠٤	٠٨	١٣	١٧	٢١	٢٦	٣١	٣٦	٤٢	٤٨	٥٦	٦٨
٥٤														٠٠	٠٤	٠٨	١٣	١٧	٢٢	٢٧	٣٢	٣٨	٤٥	٥٣	٦٦
٥٨															٠٠	٠٤	٠٩	١٣	١٨	٢٣	٢٨	٣٤	٤١	٥٠	٦٣
٦٢																٠٠	٠٤	٠٩	١٤	١٩	٢٥	٣١	٣٨	٤٧	٦١
٦٦																	٠٠	٠٤	٠٩	١٥	٢٠	٢٧	٣٤	٤٤	٥٨
٧٠																		٠٠	٠٥	١٠	١٦	٢٨	٣٠	٤٠	٥٥
٧٤																			٠٠	٠٦	١١	١٨	٢٦	٣٦	٥١
٧٨																				٠٠	٠٦	١٢	٢١	٣١	٤٨
٨٢																					٠٠	٠٧	١٥	٢٦	٤٣
٨٦																						٠٠	٠٨	١٩	٣٧
٩٠																							٠٠	١٩	
٩٤																								٠٠	
٩٨																									

وقبل ان نختتم موضوع تمييز الفقرات لا بد أن نطرح الملاحظات الآتية:

أ- عندما يتساوى عدد الطلاب الذين أجابوا على نفس الفقرة اجابة صحيحة من المجموعتين الدنيا والعليا فان معامل تمييز الفقرة يكون (صفراً) وهذا يعني ان الفقرة لم تميز في المجموعتين.

مثال:

$$0 = \frac{15 - 15}{20}$$

ب- إذا زاد عدد طلاب الفئة العليا عنهُ في الفئة الدنيا يكون معامل التمييز موجباً والعكس يؤدي الى أن يكون سالباً.

مثال:

$$0.50 = \frac{5 - 15}{20}$$

$$0.40- = \frac{15 - 7}{20}$$

جـ- قد يصل معامل تمييز الفقرة الى (1) صحيح، عندما لم يجب بشكل صحيح أي طالب من الفئة الدنيا، بينما يجيب على الفقرة بشكل صحيح جميع طلبةَ الفئة العليا.

مثال:

$$\frac{0 - 20}{20} = 1$$ وهذه حالة قليلة الحدوث.

د - وتحدث عكس هذه الحالة عندما لا يجيب أي طالب من الفئة العليا إجابة صحيحة، بينما يجيب جميع أفراد الفئة الدنيا على الفقرة إجابة صحيحة.

مثال:

$$\frac{20 - 0}{20} = 1-$$ وهذه الحالة قليلة الحدوث.

فعالية البدائل الخاطئة:

يقصد بفعالية البديل الخاطئ قدرته على جذب إنتباه الطلبة ذوي المستوى الأدنى لإختياره كبديل يمثل الإجابة الصحيحة، أما البديل الذي لا يختارهُ أي من طلاب الفئة العليا أو الدنيا فهو بديل غير فعال يُفترض حذفهُ من الاختبار، ومما تقدم نستنتج أن البديل الفاعل (أو المشتت أو المموه عن الاجابة الصحيحة كما يسمى أحياناً) هو الذي يكون مشتتاً جذاباً ومغرياً، أي إنه يبدو للطلبة الضعفاء وكأنهُ إجابة صحيحة بحيث يختاره اكثريتهم.

مثال :

يمثل الجدول الآتي نتائج إجابات طلبة الفئة العليا وطلبة الفئة الدنيا على فقرة اختبارية مفترضة، ما هو البديل الخاطىء الفعال من بين بدائل الإجابة المتوقعة لتلك الفقرة؟

البدائل / الفئات	أ	ب * الإجابة الصحيحة	جـ	د
الفئة العليا	1	12	2	–
الفئة الدنيا	9	5	1	–

من الجدول أعلاه نلاحظ أن البديل (أ) إختاره (9) من طلاب المجموعة الدنيا، وبذلك يكون هو البديل الخاطىء الفعّال، كما نلاحظ أن البديل (د) لم يختاره أي من طلاب المجموعتين مما يعني أنه بديل غير فاعل ويفترض حذفه أو إستبداله.

فوائد اخرى للتحليل الكمي لفقرات الاختبار

بالاضافة الى الفوائد السابقة الذكر التي نُجنيها من التحليل الكمي لفقرات الاختبار كإيجاد معاملات الصعوبة والتمييز وتحديد فعالية البدائل الخاطئة، فان هناك خصائص اخرى يمكن إكتشافها من خلال ذلك التحليل ومنها:

فقدان مفتاح الاجابة. **Miskeying**

وتحدث هذه الحالة عندما يختار أغلب الطلبة من الفئة العليا البديل الخادع أكثر من إختيارهم للاجابة الصحيحة.

مثال:

- من هو أول رائد فضاء نزلَ على سطح القمر؟

أ- جون كلين. ب- سكوت كاربنتر.

جـ- نيل آرمسترونك. ● د- آلن شيبرد.

وكانت اختيارات طلبة الفئة العليا من الطلبة البالغ عددهم (13) كالآتي:

إختيارات الفئة العليا

أ	ب	جـ	●د
1	1	9	2

ويلاحظ من النتائج المدرجة أعلاه ان الطلبة الاقوياء اختاروا بعدد اكبر البديل الخاطئ (جـ) مما يدل على فقدانهم لمفتاح الإجابة (د).

التّخمين Guessing

عندما يُستخدم التخمين فان طلبة الفئة العليا يختارون إجاباتهم بطريقة عشوائية بسبب أن محتوى الفقرة قد يكون:

أ- غير مطروح في الكتاب المقرر.

ب- صعب جداً حتى انهم لا يجدون له جواباً.

جـ- تافه جداً بحيث لا يعرف الطلبة أي من البدائل المعطاة يختارون.

وفي هذه الحالة تكون كل البدائل المطروحة جاذبة لطلبة هذه الفئة حتى ان اختياراتهم تتوزع بشكل متساوي بين جميع البدائل.

مثال:

أي من معاملات الارتباط الآتية يحتاجها الباحث لتحليل بيانات بحثه؟

أ- بيرسون. ب- سبيرمان. جـ - فاي. د- بايسيريال.

وفيما يلي توزيع اجابات الطلبة على البدائل المعطاة:

	أ	ب	جـ	د
اختيارات الفئة العليا	4	3	3	3

ولا شك أن السبب في هذا التوزيع يرجع الى ان طلبة الفئة العليا اعتقدوا أن اختياراتهم يمكن أن تبنى على حاجتهم الشخصية لهذه المعاملات، او نظراً لان محتوى الفقرة لم يتضمن معياراً تفصيلياً للاجابة مثل أي منها اكثر استخداماً، أو اي منها يستخدم مع متغيرات محدد نوعها وغير ذلك، مما دفع الطلبة الى اسلوب تخمين الاجابة.

الغموض Ambiguity

وتحدث هذه الحالة عندما يتجه إختيار طلبة الفئة العليا نحو أحد البدائل الخاطئة الخادعة بنفس إتجاههم نحو اختيار البديل الصحيح بحيث تتساوى الاختيارات بين هذين البديلين، ولا شك ان ذلك يُعزى الى غموض صياغة الفقرة ونقص المعلومات في أصل الفقرة التي قد تساعد الطالب على تحديد الاجابة الصحيحة.

مثال:

يطلق على الثعلب اسم إبن آوى :

أ - لأنه أصغر الحيوانات. ●ب - لأنه أذكى الحيوانات.

جـ - أطولها قامة. د - أقصرها قامة.

	أ	●ب	جـ	د
اختيارات الفئة العليا	6	6	0	1

ويلاحظ من النتائج اعلاه ان افراد الفئة العليا توزعوا بشكل متساوٍ على الاجابة الخاطئة (أ) والاجابة الصحيحة (ب) مما يعني غموض هذه الفقرة.

المعايير Norms

إن الدرجة التي يحصل عليها الطالب في اختبار ما تسمى بالدرجة الخام raw score، وهذه الدرجة لا يمكن من خلالها الحكم مباشرة على مستوى الطالب، إذ ليس لها معنى، ولا يمكن تفسيرها إلا بمقارنتها بمتوسط درجات المجموعة التي طبق عليها الاختبار، إن المتوسط الذي تتم مقارنة الدرجة الخام به يمثل الأداء الطبيعي أو العادي في المجتمع (أو الصف الدراسي للطالب). فمثلاً إذا حصل الطالب أحمد على الدرجة 48 في مادة الرياضيات، فإن هذه الدرجة لا تدل على ما إذا كان هذا الطالب جيداً في هذه المادة، أو متوسطاً، أو ضعيفاً، فقد يكون الامتحان صعباً وهذه الدرجة هي أعلى الدرجات، وقد يكون الامتحان سهلاً وهذه الدرجة هي أقل الدرجات، أو قد يكون متوسطاً وهذه الدرجة تقع وسط التوزيع. من جهة أخرى لا بد أن نعرف الدرجة القصوى في الامتحان، هل هي (100)؟ أم (50)؟ فإذا كانت (100) فإن احمد راسب في الإمتحان، وإذا كانت (50) فإن احمد من الطلبة المتازين. إن نقاط ضعف الدرجة الخام لا تقتصر فقط على تلك الجوانب، بل هناك نقاط ضعف أخرى، ومنها عدم امكانية المقارنة بين الدرجات الخام نفسها، فمثلاً درجة أحمد في الرياضيات 48 وفي العلوم 68، في هذه الحالة لا نستطيع ان نقارن بين الدرجتين، كأن نقول أن درجته في العلوم أفضل من الرياضيات؛ فربما تكون درجته في الرياضيات هي اعلى درجة بين طلبة صفه، ودرجته في العلوم هي في الوسط.

وعلى أساس ذلك لابد من مقارنة الدرجة الخام بنقطة مرجعية، ومن خلال هذه النقطة المرجعية نحكم على مستوى الطالب إن كان جيداً او ضعيفاً.

فالمعيار هو ميزان يحدد مستوى اداء مجموعة من الطلبة في امتحان ما، وتستعمل للحكم على مستوى أداء أي طالب بالنسبة لأداء المجموعة التي ينتمي اليها.

خصائص عينات اشتقاق المعايير

يتم اشتقاق المعايير عادة من عينات محددة ذات مواصفات خاصة. ويشير ويرسما وجورز (Jurs & Wiersma ،1990) الى ثلاث خصائص أو شروط لابد من توافرها في العينة التي تشتق منها المعايير وهي:

1- تمثيل العينة لأفراد المجتمع Representatives من اكبر مشكلات التقنين (اشتقاق معايير للإختبارات) مدى تمثيل العينة لجميع خصائص المجتمع ذات الصلة بموضوع الاختبار. وهذا التمثيل لا يعبر عنه حجم العينة وإن كان الحجم احد مؤشرات التمثيل، وانما من خلال اعطاء فرصة لكل خاصية في المجتمع لان تظهر في العينة.
فمثلاً اذا كان متغير الصف الدراسي والجنس مهمان في قياس الذكاء، فلا بد من إعطاء فرصة لطلبة كل صف دراسي وللذكور والإناث لأن يظهروا في العينة.

2- الحداثة Recency ان المعايير تتغير مع الزمن، لا سيما إذا كانت الخصائص المقاسة تتأثر تاثراً ملحوظاً بما يطرأ على البيئة من تطورات. فمثلاً المفردات اللغوية تنمو وتتغير دلالتها، والمناهج الدراسية تتطور. لذلك فإن معايير الاختبارات التحصيلية تتاثر بجميع هذه العوامل التي تنعكس على الطلاب، بل وربما تتأثر اختبارات الاستعدادات ومقاييس الشخصية أيضاً ببعض هذه العوامل وغيرها.

3- المواءمة Relevance تتحقق المواءمة عندما تتفق خصائص المجتمع مع خصائص العينة التي سيطبق عليها الاختبار، فإذا تم اشتقاق معايير لمجتمع طلبة مرحلة تعليمية، فلا بد أن يستخدم مع المدى العمري لتلك المرحلة وليس لأعمار اكبر او أصغر منها.

مستويات المعايير

يشير علام (2006) الى وجود اربعة مستويات للمعايير هي:

1- المعايير الوطنية National Norms

وهي معايير على مستوى انحاء الوطن الواحد، تضم عيناتها جميع المناطق والمحافظات والولايات. يتم اختيار العينة باستخدام أساليب المعاينات العشوائية او العشوائية الطبقية التي تراعي تمثيل خصائص المجتمع، ويتعين أن تكون العينة المنتقاة كبيرة بدرجة كافية لضمان ملاءمة المعايير للمجتمع المستهدف، وتقليل اخطاء المعاينات.

2- المعايير المحلية Local Norms

وهي معايير على مستوى منطقة تعليمية، أو ولاية، وفي بعض الاحيان هذه المعايير اكثر فائدة من المعايير الوطنية؛ فنادراً ما نجد تطابقاً بين خصائص الجماعة المرجعية التي تستمد منها المعايير الوطنية وخصائص الجماعة التي يطبق عليها الاختبار فيما بعد؛ إذ أن هذه المعايير يختلف معناها بمرور الزمن، وتغير خصائص الأفراد مما يؤدي الى صعوبات في تفسير الدرجات تفسيراً ذي معنى.

3- معايير مؤسسة أو مدرسة Institution Average Norms

وهي معايير مناسبة لطلبة مؤسسة تعليمية او مدرسة معينة، وتفيد في مقارنة متوسط أداء شعبة مع شعبة أخرى بنفس المستوى، فمثلاً يمكن مقارنة متوسط اداء الشعبة (أ) بالصف الخامس مع أداء الشعبة (ب) بنفس الصف. ومما تجدر الاشارة اليه ان هذه المقارنة تتم فيها مقارنة مجموعة مع مجموعة وليس طالب مع طالب.

4- معايير فئات خاصة Special Norm Groups

في بعض الحالات نحتاج الى اشتقاق معايير لمجموعة او لفئة من الطلبة، مثلاً فئات الطلبة ذوي الاعاقات العقلية او الجسمية في بلد ما، او نحتاج الى اشتقاق معايير لطلبة الصف الاول بالتعليم الأساسي في مدرسة ما. ففي مثل هذه الحالات ربما تبدو ضرورة إعداد معايير خاصة لهذه الفئات للاسترشاد بها في اتخاذ قرارات مناسبة بشأنهم.

أنواع المعايير Type of Norms

هناك أنواع عديدة من المعايير، ويمكن تصنيفها الى نوعين رئيسيين هما:

أ- **المعايير الإحصائية Statistical Norms** : وهي المعايير التي تقسم المستويات أو الفروق الفردية على المحور الأفقي للمضلع التكراري إلى وحدات قياسية متساوية، وثابتة مثل: المئينات والدرجة التائية والدرجة المعيارية ... الخ.

ب- **المعايير الوصفية Descriptive Norms** : وهي المعايير التي توضع تبعاً لمتغيرات معينة مثل: المعايير العمرية، والمعايير الصفية .. الخ.

وفي مجال القياس والتقويم التربوي نحتاج إلى المعايير الإحصائية والوصفية على حد سواء، إلا ان ما يهمنا هنا هو المعايير الإحصائية فقط، لأنها على صلة وثيقة بتفسير نتائج الاختبار.

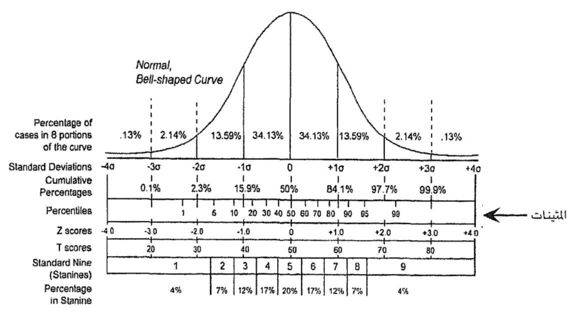
واكثر المعايير الاحصائية التي تشتق في اختبارات التحصيل هي المئينات والدرجة المعيارية. سنتناول كلاً منها من حيث المفهوم وطريقة الحساب.

1- المئينات : Percentile

المئين هو النقطة التي تقسم المحور الأفقي للمضلع أو للمنحنى التكراري إلى مائة قسم متساوٍ، وعليه فإن أول نقطة هي المئين 1 ويرمز لها م1 وآخر نقطة هي المئين 99 ويرمز لها م99 .

فالمئين الأول (م1) يُعرّف بأنه القيمة التي يكون $\frac{99}{100}$ من الدرجات أقل منها أو يساويها، ويكون $\frac{1}{100}$ من الدرجات أعلى منها أو يساويها، بافتراض أن الدرجات مرتبة تصاعدياً أو تنازلياً، والمئينات الإخرى من م^2 إلى م99 تُعرف بنفس الطريقة.

والشكل (7-1) يوضح مفهوم المئينات



الشكل (7-1) منحنى بياني يوضح المئينيات

وتختلف طريقة حساب المئين باختلاف نوع البيانات، ففي حالة البيانات الأولية يحسب المئين على وفق المعادلة الآتية:

$$\text{رتبة المئين} = \frac{\text{المئين المطلوب} \times \text{عدد الدرجات}}{100} + 1$$

مثال:

أوجد المئين 50 للدرجات الآتية:

20	15	7	19	28	10

الحل

1- نرتب الدرجات تصاعدياً.

28	20	19	15	10	7

2- نطبق معادلة المئين:

$$\text{رتبة المئين } 50 = \frac{6 \times 50}{100} + 1 = 4$$

المئين 50 هو الدرجة المقابلة للقيمة الرابعة (وهي 19).

وفي بعض الأحيان يحصل العكس، أي نريد أن نعرف المئين المقابل لدرجة معينة، وليس معرفة الدرجة المقابلة لمئين معين، في هذه الحالة نستخدم المعادلة الآتية:

$$\text{مئين الدرجة المطلوبة} = \frac{\text{عدد الدرجات الأقل من الدرجة المطلوبة}}{\text{عدد الدرجات}} \times 100$$

فلو أردنا أن نعرف مئين الدرجة 19، نرتب الدرجات الخام كما فعلنا في المثال السابق، وبعدها نطبق المعادلة الآتية:

مئين الدرجة 19= $\frac{3}{6} \times 100 = 50$ وهو المئين المقابل للدرجة 19.

واما إذا كانت البيانات موزعة تكرارياً، فنستخدم المعادلة الآتية عندما نريد أن نحسب الدرجة المقابلة لمئين معين:

$$\text{الدرجة المقابلة للمئين المطلوب} = \text{الحد الفعلي الأول لفئة المئين} + \frac{\frac{\text{المئين المطلوب}}{100} \times \text{عدد الدرجات} - \text{التكرار المتجمع السابق للمئين}}{\text{تكرار فئة المئين}} \times \text{طول الفئة}$$

$$= \text{أ} + \frac{\frac{\text{ك}}{100} \times \text{ن} - \text{ت ق}}{\text{ت}} \times \text{ف}$$

مثال:

أوجد المئين 70 والمئين 90 للتوزيع التكراري الآتي الجدول رقم (9):

الفئات	التكرار	الحدود الفعلية للفئات	التكرار المتجمع الصاعد
36-32	9	36.5-31.5	9
41-37	6	41.5-36.5	15
46-42	10	46.5-41.5	25
51-47	12	51.5-46.5	37
56-52	11	56.5-51.5	48
61-57	10	61.5-56.5	58
66-62	2	66.5-61.5	60

الحل :

1- نستحدث خانة للحدود الفعلية للفئات وخانة أخرى للتكرار المتجمع الصاعد.

2- نحدد الفئة التي يقع فيها المئين 70 = $\frac{70}{100} \times 60 = 42$ وهي الفئة الخامسة، لأنها أول فئة يزيد تكرارها المتجمع الصاعد عن 42.

3- نطبق معادلة المئين:

$$م_{70} = 51.5 + \frac{60 \times \frac{70}{100} - 37}{11} \times 5$$

= 53.77 نقربها تصبح 54 وهي الدرجة المقابلة للمئين 70.

4- لإيجاد المئين 90، نحدد الفئة التي يقع فيها $م_{90}$ = $\frac{90}{100} \times 60$ = 54 وهي الفئة السادسة، لأنها أول فئة يزيد تكرارها المتجمع الصاعد عن 54.

5- نطبق معادلة المئين:

$$م_{90} = 56.5 + \frac{\frac{90}{100} \times 60 - 48}{10} \times 5$$

= 59.5 نقربها تصبح 60 وهي الدرجة المقابلة للمئين 90.

2- الدرجة المعيارية: **Standard Score**

وهي المسافة التي تبعدها الدرجة الخام عن المتوسط الحسابي، معبراً عنها في وحدات من الإنحراف المعياري.

وتعبر الدرجة المعيارية عن موقع أي درجة خام في قربها أو بعدها عن المتوسط، حيث تقسم قاعدة التوزيع إلى أقسام قياسية متساوية، وتعد كل واحدة إنحرافاً معيارياً واحداً.

والدرجة المعيارية أكثر دقة من المئين، لأن المئين يتأثر بالدرجات المتطرفة للتوزيع، بينما الدرجة المعيارية لا تتأثر بذلك، ويمكن الرجوع إلى الشكل (7-2) لمعرفة موقع الدرجات المعيارية مقارنة بالمئين وغيره من المعايير.

ويعتمد حساب الدرجة المعيارية على الوسط الحسابي، والإنحراف المعياري على وفق المعادلة الآتية:

$$\text{الدرجة المعيارية} = \frac{\text{الدرجة الخام} - \text{الوسط الحسابي}}{\text{الإنحراف المعياري}}$$

$$ز = \frac{س - \overline{س}}{ع}$$

مثال:

أوجد الدرجات المعيارية للدرجات الخام (35 ,42 ,53) إذا علمت أن الوسط الحسابي 40 والإنحراف المعياري 5 .

$$ز_1 = \frac{35 - 40}{5} = -1$$

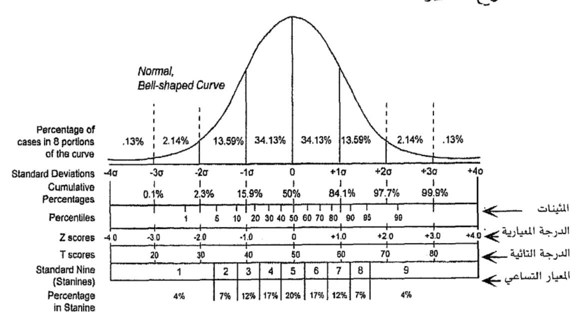
$$ز_2 = \frac{40 - 42}{5} = 0.4$$

$$ز_3 = \frac{40 - 53}{5} = 2.6$$

لاحظ بأن الدرجة المعيارية الأولى سالبة، والثانية والثالثة موجبة، والإشارة السالبة تعني أن الدرجة أقل من الوسط الحسابي، والموجبة تعني أنها أعلى من المتوسط.

تحويل المعايير

إن تعدد أنواع المعايير لا يعبر عن اختلافها وتباينها، فهي وجوه لعملة واحدة، وهناك امكانية او قابلية لتحويلها من نوع لآخر Interchange ability of Norms وهذه الخاصية المهمة التي تتمتع بها المعايير الاحصائية تتيح للباحث امكانية المقارنة بين المعايير المئينية والدرجات المعيارية مثلاً، او بين الدرجات المعيارية والدرجات التائية وهكذا. والشكل البياني (7-2) يوضح العلاقة بين مختلف انواع المعايير.



الشكل (7-2) : منحنى بياني يوضح المعايير الاحصائية

فمثلاً لو كان توزيع درجات الطلبة إعتدالي، وحسبنا المئين (85)، نستطيع

معرفة المعايير الأخرى بالاستفادة من خاصية تحويل المعايير دون الحاجة لتطبيق المعادلات الإحصائية الخاصة بها. فمن خلال الشكل البياني (7-2) نحدد المئين (85)، وبعدها نضع بالمسطرة خط عمودي لنجد أن الدرجة المعيارية المقابلة له هي (1+)، وأن الدرجة التائية المقابلة له هي (60)، وأن المعيار التساعي المقابل له هو (7). وهكذا يمكن الإستفادة من خاصية تحويل المعايير في حالات عديدة مشابهة.

خصائص المعايير ومحاذير استخدامها

يرى ثورندايك وهيجن (1989) إن المبالغة في استخدام المعايير والاعتماد عليها بشكل مطلف يمكن ان يؤدي الى نتائج سلبية ,وذلك يرجع الى الاسباب الآتية:

1- ان المعلومات التي تقدمها المعايير هي معلومات نسبية وليست مطلقة، فهي تبين ما إذا كانت درجات صف معين بمستوى الصفوف الاخرى، ولكنها لا تبين ما إذا كانت القدرة او الظاهرة المقاسة قد وصلت الى مستوى الاتقان والتمكن، فعندما يكون مئين الطالب أحمد (95) في مادة اللغة الانجليزية بالصف الثالث أساسي، فهذا لا يعني أنه قادر على كتابة رسالة باللغة الانجليزية الى معلمته يخبرها فينا بانه يحب طريقتها في التدريس، وذلك لأن كتابه الرسالة غير مرتبط باهداف تدريس اللغة الانجليزية بالصف الثالث الأساسي، وهكذا مع باقي المقررات الدراسية.
2- يجب ان تقوّم المخرجات بدلالة المدخلات؛ فنتائج الاختبارات تعطينا صورة عن المخرجات، أي عن مستوى الفرد او المجموعة كما هي في الوقت الحاضر. وذلك بسبب تعقد مفهوم المدخلات، فلا يمكن قياس المدخلات جميعها، فبعضها يرجع للظروف الاسرية والمؤثرات البيئية التي تجعل تعلم بعض التلاميذ اسهل من تعلم بعضهم الاخر.
3- يجب ان تقوّم المخرجات بدلالة الأهداف؛ فمحتوى الاختبارات ومعاييرها تستند بالدرجة الاساس الى تصورات مؤلفي هذه الاختبارات. ان تحديد اهداف للمنهج لا يعني تشابه المحتوى وتساوي المعايير عندما يتناولها أكثر من مؤلف، وهكذا فإن دور المؤلف (سواء كان المعلم او لجنة متخصصة) كبير جداً في تحديد نوع المحتوى، ومستوى صعوبة الاختبار، وبالتالي مستوى المعايير التي يمكن ان نقيس في ضوئها أداء الطالب.

خطوات بناء الاختبارات التحصيلية

من المعروف أن التخطيط السليم والدقيق يؤدي الى نتائج أكثر دقةً وموضوعية، وحيث أننا نستهدف في الاختبار التحصيلي أن يضمن قياساً واقعياً لأداء الطالب المعرفي، المتمثل في عينة ممثلة من سلوكه، وما ينشأ عن هذه المهمة من إتخاذ قرارات قريبة أو بعيدة تتعلق بالطالب أو باداء المعلم أو تقويم المنهج الدراسي. ومن هنا أصبحت عملية بناء الاختبارات التحصيلية تتطلب أمرين هامين هما:

1 - مهارة المعلم في بناء الاختبارات:

ويتطلب هذا الأمر أن يكون المعلم ذو خبرة وكفاءة في بناء الاختيارات، هذه الخبرة اكتسبها من خلال دراسته في مؤسسات إعداده لمقررات تركز على هذا الجانب، أو من خلال الورش التدريبية التي ينخرط فيها أثناء خدمته الوظيفية. فاذا ما تعززت هذه الخبرات بالواقع العملي ومتابعة المصادر العلمية، فقد يصل المعلم الى المستوى الذي يمكن أن يساعده على صياغة فقرات اختبار موضوعي أو مقالي جيد.

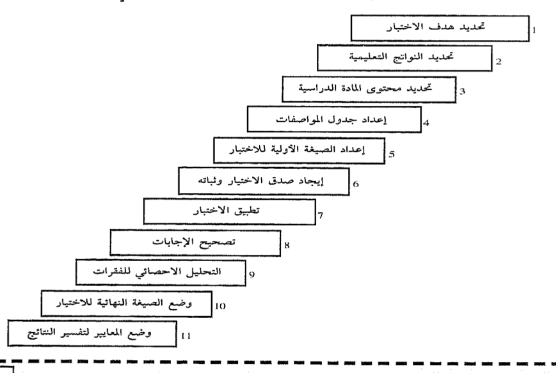
2 - الاستخدام الأمثل للنتائج:

وهذه الخطوة تكمل الخطوة السابقة، فالمعروف أن كل أداة قياس يجب أن تستخدم بشكل سليم في الموقف الذي أُعدَت من أجله. ومهما يكن الاختبار مصمماً بشكل جيد ، فلا فائدة ترجى منهُ اذا لم يستخَدم استخداماً صحيحاً، فليس صحيحاً مثلاً أن تستخدم اختبار وُضِعَ لقياس القدرة الحسابية لطلاب الصف السادس الابتدائي في قياس تحَصيل طلبة الصف الخامس في الرياضيات، اذ يفترض أن يقيس الاختبار ما وضعَ لقياسه لدى المجموعة المستهدفة عينها وليس مجموعات اخرى. كما أن النتائج المتمَخضة عن تطبيق الاختبار يجب أن تستخدم في حّدود الاهداف المحددة لاستخدامها ولا تعمم إلا في حدود تلك المجموعة لأنها تمثل ادائهم وليس أداء مجموعات اخرى، وان استخدامها أيضاً مطلوب لاتخاذ قرارات تتعلق بمستقبل تلك المجموعة فقط. وبالاضافة إلى ما ذُكر فان هناك فوائد أخرى يجنيها المعلم من عملية تخطيط وبناء الاختبار الجيد منها إمكانية الاحتفاظ بالفقرات الجيدة للاختبار

والتي تُشير التحليلات الاحصائية الى كفائتها وتجميعها بشكل تراكمي مستمر لتكوين ما يسمى ببنك الأسئلة Item Bank حيث يستفيد منها في اعداد اختبارات قادمة لطلبته مما يخفف عنه مهمة الجهد المتكرر كلما اراد اختبار طلبته، اذ يصبح بامكانه الاستعانة بتلك الفقرات لا سيما وأن الاجهزة الالكترونية كالحاسوب أصبح يسهل عمليات التحليل والخزن والاعداد بأقل وقت وجهد.

لقد ركزنا في الفصول السابقة من هذا الباب على مواصفات الاختبار الجيد وشروط تطبيقه وتصحيحه وفقاً للقواعد العلمية، وتطرقنا للكثير من الامور الهامة لبناء الاختبار التحصيلي وهي جميعها مهمة جداً لعملية البناء، وتعتبر خطوات ضرورية لتلك العملية، وبذلك أصبح بمقدور المعلم أو الطالب الجامعي خلال دراسته لمقرر القياس والتقويم التربوي ان يصمم اختباراً تحصيلياً ذو مواصفات عالية، لذلك فاننا سوف لن نكرر ذلك في عملية التخطيط والبناء، بل سنقدم فيما يأتي تسلسل تلك الخطوات وبامكان المعلم او الطالب مراجعة الفصول السابقة للاستفادة منها في عملية البناء. تمر عملية بناء الاختبار التحصيلي بعدد من الخطوات يمكن تمثيلها في الشكل (7-3):

الشكل (7-3) : خطوات بناء الاختبار التحصيلي



مراجع الفصل السابع

1- ثورندايك، روبرت، وهيجين، اليزابيث هيجن (1986). **القياس والتقويم في علم النفس والتربية**، ترجمة عبد الله الكناني وعبد الرحمن عدس. الأردن: مركز الكتاب الأردني.

2- ديرجي، نجيب أحمد؛ والألوسي، جمال حسين (1980). موضوعية تقويم الاجابات الامتحانية في المرحلة الثانوية. **مجلة العلوم التربوية والنفسية** (5). بغداد.

3- العاني، نزار محمد سعيد (2002). **القياس والتقويم وبناء الاختبارات المدرسية**. الكويت: الجامعة العربية المفتوحة.

4- عبد الرحمن، سعد (2003). **القياس النفسي النظرية والتطبيق**. القاهرة: دار الفكر العربي.

5- عدس، عبد الرحمن (1999). **دليل المعلم في بناء الاختبارات التحصيلية**. عَمّان: دار الفكر للنشر والتوزيع.

6- علام، صلاح الدين محمود (2006). **القياس والتقويم التربوي والنفسي: أساسياته وتطبيقاته وتوجهاته المعاصرة**. القاهره: دار الفكر العربي.

7- عيد، غادة خالد (2007). **القياس والتقويم التربوي مع تطبيقات برنامج SPSS**. الكويت: مكتبة الفلاح للنشر والتوزيع.

8- كاظم، علي مهدي (2001). **القياس والتقويم في التعلم والتعليم**. عَمّان: دار الكندي للنشر والتوزيع.

9- مراد، صلاح؛ وسليمان، أمين (2002). **الاختبارات والمقاييس في العلوم النفسية والتربوية: خطوات اعدادها وخصائصها**. الكويت: دار الكتاب الحديث.

10- Batenburg. T. & Laros, J. (2002) Graphical qnalysis of test items. **Educational Research and Evaluation,** 8 (3), 319-333.

11- Chase, I. C. (1978). **Measurement for Educational Evaluaation (2n ed).** Blamingtion, Indiana: Addison- Wesley pulishing Company.

12- Eble, R. (1975). Can teachers write good true- false test items? **Journal of Educational Measurment**, 12: 31-36.

13- Eble, R. (1979). **Essentials of Educational Measurement (3rd ed.).** New Jersey: Prentic- Hall.

14- Frisbie, D. A. (1974) The effect of them formate on reliabilty and validity: A study of multiple- choice and true- false achievement tests. **Educational and Psycholog,cal Meccsurement** 34, 885-892.

15- Gronland, N. (1985). **Measurement and Evaluation teaching (5th ed.).** New york: Macmilan.

16- Gronland, N. (2006). **Assessment of Student Achievement**, Boston: Person Educatin, Inc.

17- Halagyna, T. M., & Downing, S. M. (1993). How many options is enough for a multiple - choice test item. **Educational and Psychological Measurement**, 53, 999-1010.

18- Haladyna, T. M. (1997). **Writing Test Items to Evaluate Higher Order Thinking. Need ham Heights**, MA: Allyn and Bacon.

19- Haladyna, T. M. (2004). **Developing and Validating Multiple - choice test Items. (3rd ed.).** NJ: Lawrence Erlbaum Associates.

20- Kubiszyn, T., & Borch, G (1987). **Educational Testing and Measurement (2nd. ed).** Illinois: Scott, Foresman and Company.

21- Lindvall, C.M. (1968) **Testing and Evaluation in Education**, New York.: Inc Company

22- Mehrens, W. & Lehmann, I. (1991). **Measurement and Evaluation in Education and Psyehology**. New York: Harcourt Brace College Publishers.

23- Oosterhof, A. (2001), **Classroom Applications of Educational Measurement.** New Jersy: Merrill Preutice - Hall.

24- Osterland, S. J. (2001). **Constructing Test Items: Multiple Choice, Contstructed - response, Perfrmance and Other Formats (2nd ed).** Boston: Kluwer Academic Publi - shers.

25- Pohham, W. (1981). **Modern Educaational Measurement**. Englewood Cliffs, New Jersey: Prentice- Hall.

26- Trice, A. (2000) **A Hanal book of Calssroom Assessment.** New York: Longman.

27- Wiersma, W. & Jurs, S. (1990). **Educational Measurement and Testing.** Massachustts: Allyn and Baconlnc.

الفصل الثامن

التقويم البديل لاختبارات الورقة والقلم

اختبارات الأداء

الملف الوثائقي

8

الفصل الثامن

التقويم البديل لاختبارات الورقة والقلم

Alternative Assessment

يعتقد الكثير من الناس بأن القياس والتقويم التربوي هو مجموعة من أسئلة الاختيار من متعدد أو الصواب والخطأ أو الأسئلة المفتوحة وغيرها من أسئلة الورقة والقلم التي عادة ما تعطى للطالب مرة أو عدة مرات في السنة بغرض الحصول على معلومات (درجات) سهلة التبليغ إلى السلطات التربوية وأولياء الأمور والطلاب، أنفسهم، وكأنها الوسيلة الوحيدة والفاعلة لقياس تحصيل الطلاب في حين أن هذه الوسيلة بالرغم من بعض ميزاتها إلا أن الدراسات والبحوث في مجال الاختبارات الصفية أشارت إلى العديد من المشكلات وأوجه القصور التي تكتنفها (السعدي، Alsarimi, 2000, 2001)

لقد عرفت اختبارات الورقة والقلم بأنها النموذج القديم أو التقليدي في القياس والتقويم التربوي والنفسي وأصبح يطلق عليها فعلا الاختبارات التقليدية (Lane, Parke, & Stone, 1999, 2002)، وقد هيمن هذا النموذج من التقويم في التعليم لقرون عدة، ومن خصائص هذا النموذج اعتماده الكبير على الأسئلة المقالية والموضوعية وخاصة أسئلة الاختيار من متعدد multiple-choice items كما أنه قليل التكلفة، ومناسب لتقويم مجموعات كبيرة من الطلاب في زمن قصير جداً، حيث كان ينظر الى هذا النوع من الاختبارات حسب الدوسري (2004) بأنه عادل؛ لأنه يعطي فرصاً متساوية في الأداء لجميع الطلبة.

وما زالت اختبارات الورقة والقلم تفرض نفسها بقوة ليس في الغرب فحسب، بل وفي معظم دول العالم، ولكنها بدأت في التراجع بدءاً من مطلع العقد الثامن من القرن العشرين، فقد بينت الدراسات في دول عديدة من العالم بأن هذه الاختبارات مقتصرة على قياس القدرات العقلية الدنيا للطالب

بدلاً من التركيز على قدرات عقلية عليا أكثر فائدة وإلحاحاً (Alsarim, 2002 Brookhart, 1993, 2001) كما أن تسارع الثورة الصناعية في المجتمعات الغربية، ومشروع غزو الفضاء الذي بدأه الاتحاد السوفيتي سابقاً، والثورة التكنولوجية والمعرفية التي تميز بها القرن العشرين شكلت ضغوطاً متواصلة على المؤسسات التعليمية لتحسن من نواتجها، ويعتبر تقرير "أمة في خطر" الذي صدر في الولايات المتحدة عقب إعلان الاتحاد السوفيتي إنزاله أول رائد فضاء على سطح القمر، والذي يصف أوجه العجز في النظام التعليمي الأمريكي الذي جعل أمريكا تتخلف عن الاتحاد السوفيتي، يعتبر بمثابة ناقوس الخطر الذي دُق ليأذن بتغيرات جوهرية في المنظومة التعليمية ككل ومن بينها أساليب التقويم المختلفة. فبدأت في أواخر الستينات من القرن نفسه ظهور مصطلحات لم تكن مألوفة في القياس التربوي مثل التقييم التربوي educational assessment، والتقويم المعتمد على الأداء performance-based assessment، والتقويم الحقيقي أو الأصيل authentic assessment، والملف الوثائقي أو الصحيفة الوثائقية portfolio.

كما أن مصطلح المساءلة accountability في النظم التربوية بدأ في الظهور في الأدبيات التربوية في أواخر الستينات من القرن العشرين، وشاع أكثر في نهاية الثمانينات من القرن نفسه (Stiggins,1991). والمساءلة كانت نتاج الشك الذي بدأ يساور الرأي العام في الغرب حول جدوى ما يتعلمه التلاميذ في مؤسسات التعليم العام، فلم يعد مقبولاً لدى عامة الناس بأن تتحمل المدرسة مسؤولية توفير فرص التعلم للتلاميذ، وإنما في الوقت نفسه يجب عليها أن تتحمل ماهية ونوعية هذا التعلم، فأصبحت المدارس مسؤولة أمام الهيئات الرسمية عن التعليم في تلك الدول عن مستوى تحصيل تلاميذها، حيث عادة ما تقوم تلك الهيئات في بداية كل سنة دراسية بتحديد مستوى تحصيلي معين يجب على المدرسة تحقيقه بنهاية العام الدراسي، فإذا ما استطاعت المدرسة تحقيق ذلك المستوى فستحصل على مكافآت مالية مجزية، وإذا ما فشلت في تحقيق ذلك فإنها تكون عرضة للحرمان من مثل هذه المكافآت ولربما يصل الأمر

إلى استقطاع جزءا من ميزانيتها السنوية، ونظام المساءلة برغم بعض المساوئ المصاحبة له، والمتمثلة في دفع المعلمين للتدريس من أجل الاختبار، إلا أنه أفرز العديد من الممارسات الجيدة في مدارس التعليم العام ولعل أهمها:

1- حرص المدارس على تحديد أهداف واضحة ومحددة لها، تكون معلنة للتلاميذ والمعلمين وأولياء الأمور، وتحدد سير العمل في المدرسة.
2- تبني المدارس نظم إدارية أكثر كفاءة وفاعلية.
3- والتركيز على قياس القدرات العقلية العليا في الاختبارات، وتنويع وسائل قياس تحصيل الطلاب.

كل ذلك أثر بشكل مباشر على نوعية التقويم المستخدم في المؤسسات التعليمية، حيث لم يعد مقبولاً أن تركز هذه المدارس على اختبارات الورقة والقلم فقط لقياس تحصيل التلاميذ، خاصة وأن هذه الاختبارات عادة ما تكون من نوع الاختيار من متعدد، وتقيس قدرات عقلية دنيا كالتذكر والفهم، كما أنها تقيس فقط الجوانب المعرفية للطالب، وتهمل الجانب المهاري، وكلنا يعرف بأن معرفة الطالب بالشيء لا يعني بالضرورة بأنه قادر على تطبيقه في الواقع، فيستطيع الطالب بدقة متناهية وصف خطوات القيام بتجربة في الكيمياء، ولكنه ليس بالضرورة قادراً على إجراء تلك التجربة في المعمل والتوصل إلى نتائج مقبولة. ومن هنا ظهرت الحاجة إلى ما يعرف بأساليب التقويم البديل للورقة والقلم alternative assessment.

ويقصد بالتقويم البديل انه الاسلوب الذي يعطى من خلاله الطلاب نشاطات ومواقف تعليمية، ويكلفون بأداء مهام وتكليفات تتشابه الى حد كبير مع مواقف الحياة الحقيقية، ومن ثم فإن ما يتم تقويمه هو أداءات حقيقية واقعية ترتبط بحياة الطلاب وواقعهم وليس مجرد استرجاع حقائق ومعلومات منعزلة عنها، يتم تلقينها داخل غرفة الدراسة (عيد، 2006).

ومن هذه الاختبارات اختبارات الأداء والملف الوثائقي، وسوف نستعرض في هذا الفصل هذين النوعين من الاختبارات البديلة، ماهيتها، وكيفية إجرائها، ومميزات كل نوع وعيوبه.

اختبارات الأداء Performance assessment

ما المقصود باختبارات الأداء؟ يعرف نيتكو وبروكهارت (Nitko & Brook-hart, 2007) اختبارات الأداء بأنها إجراء تستخدم فيه المهام للحصول على معلومات عن مدى جودة تعلم الطالب وقدرته على تطبيق ما تعلمه من معرفة ومهارات في مواقف جديدة. أما بروالدي (Brualdi, 1998) فيعرفها بأنها مجموعة من الاستراتيجيات لتطبيق المعرفة والمهارات وعادات العمل من خلال أداء المتعلم لمهام محددة، ينفذها بشكل عملي، وهي مرتبطة بواقع الحياة وذات معنى بالنسبة له. أي أن اختبارات الأداء عبارة عن مهام معينة يتم من خلالها ملاحظة سلوك الطالب للتحقق من تمكنه من تطبيق المعارف والمهارات التي تعلمها في مادة معينة أو مجموعة من المواد في مواقف جديدة بالنسبة له، ويتم ذلك من خلال قياس أو ملاحظة الإجراءات procedures أو الطريقة process، التي يتبعها الطالب في أداء المهمة كأن نلاحظ الخطوات التي يقوم بها التلميذ أثناء رسمه للوحة معينة، أو قياس المنتج النهائي product، كأن نقوم بتقييم اللوحة بعد رسمها، كما يمكن قياس الاثنان معاً أي الإجراءات والمنتج.

إن **قياس الإجراءات أو العملية** في اختبارات الاداء يكون ممكناً في الحالات الآتية:

أ- الخطوات أو الإجراءات واضحة ومحددة ويمكن التعرف عليها كل على حدة.

ب- انحراف الطالب عن المسار الصحيح لتسلسل الخطوات المختلفة يسهل التعرف عليه وقياسه.

ج- كل أو معظم الأدلة على تحصيل الطالب للمهارة يمكن الحصول عليها من الإجراءات أو العملية فقط، ولا توجد أدلة لا يمكن الحصول عليها إلا بقياس المنتج.

د- توفر الوقت والمصادر المالية والبشرية اللازمة لملاحظة وتدوين وتقييم أداء الطلاب أثناء قيامهم بالعمل، حيث أن ذلك يتطلب وجود عدد كبير من المراقبين خاصة إذا كانت المهارة موضع القياس دقيقة جداً.

أما قياس **المنتج النهائي** فيمكن في الحالات الآتية:

أ- كل أو معظم الأدلة على تحصيل الطالب للمهارة يمكن الحصول عليها من المنتج ولا توجد أدلة لا يمكن الحصول عليها إلا بقياس الإجراءات أو العملية.

ب- إمكانية الحصول على نفس المنتج باستخدام طرق أو إجراءات مختلفة، أو أن ترتيب تلك الطرق غير واضح ويصعب قياسه.

ونلجأ إلى قياس الإجراءات والمنتج معاً إذا كانت الأدلة على تحصيل الطالب للمهارة متوافرة في الطريقة والمنتج معاً، فهناك ثمة أدلة على تحصيل الطالب للمهارة لا يمكن الحصول عليها إلا من خلال ملاحظة الإجراءات، وأدلة أخرى لا يمكن التأكد منها إلا من خلال تقييمنا للمنتج، وهذا الخيار الأكثر تكلفة في اختبارات الأداء لأنه يستغرق وقتاً وجهداً أكبر، ويحتاج بالتالي إلى موارد مالية كبيرة.

هذا وتوجد مجموعة من العوامل ساهمت في شيوع هذا النوع من الاختبارات منها:

1- إقتصار اختبارات الورقة والقلم على قياس معلومات الطالب في المهارات المقاسة وليس قدرته على أداء المهارة، فالطالب قد تكون معلوماته جيدة حول المهارة ولكن ليس بالضرورة قادراً على القيام بها.

2- الطلب المتزايد على أن تكون مخرجات التعليم العام أكثر قدرة على حل المشكلات والتفكير العلمي والناقد Problem solving skills.

3- اختبارات الأداء أصبحت جزءاً من الاختبارات العامة على مستوى الدولة في الدول الغربية وخاصة الولايات المتحدة الأمريكية.

4- يعول على اختبارات الأداء أن تكون بديلاً ناجحاً ومناسباً للطلاب الذين لا يحسنون صنعاً في اختبارات الورقة والقلم.

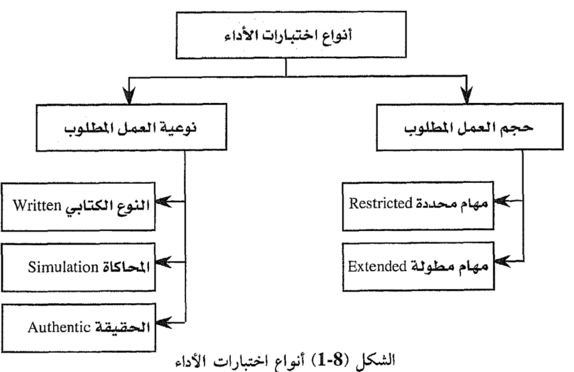
وتتكون اختبارات الأداء من جزأين رئيسين:

أ- المهمة Task وهي نشاط يهدف إلى قياس تحصيل الطالب ومدى تمكنه من أداء مهارة معينة، أي أن المهمة تعرف الطالب بما هو مطلوب القيام به، وتحدد شروط وضوابط العمل للطالب.

ب- خطة أو قواعد التصحيح Scoring rubrics وهي مجموعة من القواعد أو القوانين أو المعايير أو الضوابط مقسمة إلى مستويات حسب درجة دقتها أو صحتها أو قربها من الهدف الذي تقيسه المهمة، بحيث يُعطى كل مستوى درجة معينة، وعادة ما تكون تلك المستويات بين 3 و 7 مستويات، تستخدم في الحكم على جودة أداء الطلاب، وتعمل كمرشد للمعلم أثناء عملية التصحيح لضمان استخدام نفس المعايير أو الضوابط في تصحيح كافة الأوراق، وخطة التصحيح من الممكن أن تكون على هيئة قوائم شطب checking list أو مقياس تقدير rating-scale.

أنواع اختبارات الأداء

هناك العديد من التصنيفات لاختبارات الأداء، فالبعض يصنفها على أساس حجم العمل الذي تتطلبه المهمة (محددة أو مطولة restricted vs. ex- tended) (McMillan, 1997)، وهناك من يرى تقسيمها حسب نوعية العمل المطلوب إنجازه: اختبارات الأداء من النوع الكتابي، اختبارات المحاكاة واختبارات الأداء الحقيقية.



الشكل (8-1) أنواع اختبارات الأداء

وفيما يلي عرض موجز لكل نوع من الأنواع السابقة الذكر، وكما يأتي:

أولاً: اختبارات الأداء حسب حجم العمل المطلوب إنجازه

1- **اختبارات الأداء المحددة:** وهي عادة ما تقيس مهارات ضيقة ومحددة، ومن الأمثلة على ذلك:
   - أن يقوم الطالب بتمثيل بيانات تعطى له بيانيا.
   - التحدث حول موضوع معين بلغة أجنبية
   - قراءة موضوع معين ونقد ما جاء فيه.

2- **اختبارات الأداء المطولة:** وتتميز بأنها أكثر تعقيداً من اختبارات الأداء المحددة، حيث أنها تتطلب مدة زمنية أطول لتنفيذها وجهداً أكبر للقيام بها، وتستدعي من الطالب جمع معلومات وبيانات من مصادر مختلفة، والقيام بزيارات ميدانية، عادة ما يستغرق تنفيذها أسبوعين إلى فصل دراسي كامل، ويمكن أن تنفذ في مجموعات صغيرة من الطلاب (2 إلى 4 طلاب)، ومن الأمثلة على هذا النوع:
   - تصميم دور لرعاية الأيتام وتقدير تكلفة بنائه.
   - التخطيط لرحلة لدولة اجنبية شاملة التكاليف والأماكن التي سوف يتم زيارتها، مع تقديم المبررات.
   - عمل خطة عمل للتخلص من النفايات
   - عمل دراسة لمعالجة مشكلة اجتماعية أو اقتصادية في المجتمع المحلي.

ثانياً: اختبارات الأداء حسب نوعية العمل المطلوب إنجازه

1- **النوع الكتابي:** وهي مهام كتابية مطولة، شبيهة بأسئلة المقال المطولة، ولكنها أكثر تعمقاً، كأن نطلب من الطالب تصميم تجربة للإجابة على تساؤل معين، ففي هذا النوع من اختبارات الأداء لا يطلب من الطالب القيام بتنفيذ العمل أو المهارة، أكثر من كتابة تقرير مفصل حول خطة العمل وآلية تنفيذها.

ويمكن ان يستخدم الاختبار الكتابي كخطوة مبدئية على طريق بناء اختبار الأداء الفعلي، فإذا كان الاداء معقداً جداً، وكانت الاجهزة الواجب استخدامها ثمينة، فإن الاختبار الكتابي يمكن أن يكون خطوة مبدئية يجنبنا

مخاطر الاستخدام الفعلي للاجهزة وتعريضها للخراب أو التلف، وبالمثل في حالة العلوم الطبية فإن المهارة في تشخيص الامراض ووصف الادوية في حالات مرضية افتراضية يجنبنا الحاق الأذى بالمرضى في حالة كون التشخيص غير دقيق أو كون الدواء غير مناسب (كاظم، 2001).

مثال على اختبارات الأداء من النوع الكتابي

● جميلة تقوم بقيادة سيارتها صباحاً ومساءً كل يوم من أيام الصيف الحارة، وتريد وقاية جلدها (يديها) من آثار أشعة الشمس الضارة، المطلوب منك أن تقوم بتصميم تجربة من الممكن أن تساعد جميلة لتقرر أي نوع من أنواع الوقاية ضد أشعة الشمس سوف تستخدم.

خطة التصحيح:

الدرجة	معايير الأداء
4	يصمم الطالب تجربة تتحكم وتضبط جميع المتغيرات (كمية أشعة الشمس، ونوعية أشعة الشمس، ومعدل الرطوبة)، وتتكون التجربة من عدة بدائل أو محاولات (متغير أسلوب الوقاية المستخدم) مختلفة، وتشتمل على معايير محددة لجمع البيانات.
3	يصمم الطالب تجربة غير واضحة في كيفية التحكم في المتغيرات المختلفة ولكنها تتكون من عدة بدائل أو محاولات (أو تتحكم التجربة في المتغيرات ولكن تنقصها المحاولات) كما تشتمل التجربة على معايير محددة لجمع البيانات.
2	يصمم الطالب تجربة قد تفضي إلى نتائج وقد لا تفضي إلى نتائج وغير مكونة من عدة محاولات، ولكنها تشتمل على معايير محددة لجمع البيانات.
1	تصميم التجربة غير قائم على المنهج العلمي (يعتمد على محاولة واحدة فقط، وعدم وجود طريقة محددة لجمع البيانات) والتجربة لا تبدو بأنها ستتوصل إلى نتائج مفيدة.
صفر	الإجابة خاطئة تماماً وغير متعلقة بالموضوع

2- **اختبارات المحاكاة** Simulation ويتم في هذا النوع من الاختبارات خلق بيئة تحاكي الواقع وتكون مشابهة له بل ومطابقة له إلى حد كبير، كأن يختبر طلاب كلية الطب عملياً في مقرر التشريح باستخدام نماذج بلاستيكية لجسم الإنسان لتعرف أجزاء الجسم ووظائفه الفسيولوجية، أو كأن يتم اختبار طلاب كلية الهندسة باستخدام نموذج بلاستيكي للطائرة، أو نموذج بلاستيكي لأحد الأجهزة التي لا يمكن الحصول عليها لغلاء ثمنها أو لندرتها، وفي كليات التربية يخضع الطالب قبل أدائه التطبيق العملي في مدرسة معينة لاختبار يجرى في معامل التدريس المصغر Micro-teaching والتي هي في الحقيقة محاكاة للصف الحقيقي في المدرسة، تختبر فيه قدرات الطالب التدريسية. او اختبارات الطيارين على نماذج من الطائرات تعمل بنفس مواصفات الطائرات الاعتيادية دون ان تحلق في الجو، وكذلك في مجال التربية الرياضية كأن يطلب من الطالب التلويح بالمضرب تجاه كرة متخيلة في الهواء او الملاكمة في الهواء أمام مرآة، او توضيح حركات السباحة المختلفة خارج الماء.

3- **اختبارات الأداء الحقيقي**: وهذا ما يعرف بالتقويم الأصيل authentic assessment، ويتم فيها اختبار الطالب في بيئة العمل الحقيقية أو بيئة قريبة منها، كأن نُقِّوم طالب كلية الطب وهو يفحص مريض في إحدى العيادات، ونقوّم طالب كلية الهندسة وهو يصلح جهاز تلفزيون به عطل معين، وطالب كلية التربية أثناء إلقائه لدرس في أحد الصفوف في التربية العملية.

مثال على اختبارات الأداء الحقيقية

الهدف السلوكي:

- أن تقوم الطالبة بإعداد كعكة البرتقال.

السؤال: -

سوف تستقبل أسرتك مجموعة من الضيوف القادمين لتهنئتكم بالعيد السعيد، وطلبت أمك منك بأن تقومي بإعداد كعكة البرتقال، فبالاستعانة بالمقادير أدناه، قومي بعمل الكعكة على أن تراعي بأن تكون ذات لون ذهبي ورائحة زكية وطعم لذيذ.

المقادير:

- 2 كوب دقيق يضاف إليه 2 ملعقة صغيرة باكينج باودر
- 1 كوب سكر
- 4/3 كوب سمن
- 3 بيضات
- 4/3 كوب عصير البرتقال

خطة التصحيح:

الدرجة	معايير الأداء
4	خطوات عمل الكعكة كانت سليمة والمقادير صحيحة. والكعكة جاءت مرتفعة الحجم، وذهبية اللون، وطعمها لذيذ، وذات رائحة زكية.
3	أخطأت الطالبة في أحد المقادير، أو إجراءات التنفيذ، مما أدى إلى فقدان الكعكة لونها الذهبي وشكلها الطبيعي.
2	أخطأت الطالبة في اثنين من المقادير، ولم تتبع كافة إجراءات التنفيذ مما أدى إلى عدم نضج الكعكة من الداخل مما أثر على طعمها ولونها ورائحتها.
1	بذلت الطالبة مجهوداً واضحاً في إعداد الكعكة، ولكن محاولتها كانت فاشلة، فلم تنجح في إعداد الكعكة.

خطوات بناء اختبارات الأداء

لا تختلف خطوات بناء اختبارات الأداء عن خطوات بناء أي أداة قياس، فهي تتكون من ثلاث خطوات: التخطيط للاختبار، وكتابة مهام الاختبار، وبناء معايير أو خطة التصحيح scoring rubrics، ومن ثم تصحيح الاختبار، وتفسير الدرجات.

أ- التخطيط للاختبار:

عند بناء اختبار الورقة والقلم فإننا نقوم بتصميم جدول مواصفات، نحدد فيه موضوعات المحتوى ومستويات الأهداف المختلفة موضع القياس، والهدف من ذلك ضمان تمثيل الاختبار لجميع الموضوعات والأهداف، وفي حقيقة الأمر فإن جدول المواصفات يمكن أن ينطبق أيضاً على اختبارات الأداء مع مراعاة نوعية الأهداف، فاختبارات الأداء تركز على الأهداف النفسحركية والمهارية، إذاً فالتخطيط لاختبارات الأداء يتضمن تحديد الهدف من الاختبار، والمعلومات والمهارات والقدرات التي ننوي قياسها، وكذلك معايير الأداء الجيد في المهمة موضع القياس، كما أنه وبناء على نوعية الأداء المنصوص عليها في الأهداف، نقوم بتحدد ما الذي سنقيسه: الإجراءات أم المنتج أم الاثنين معاً، كل ذلك يجب أن يتم مع وضع قدرات الطلاب ومستوى تحصيلهم في الحسبان.

ب- كتابة مهام أو أسئلة اختبارات الأداء:

بعد تحديد المهارات التي سيتم قياسها ومعايير الأداء الجيد، نبدأ بكتابة مهام الاختبار، والتي تحدد ما يجب على الطالب عمله بوضوح ودقة، بحيث يستطيع الطالب التعرف على المطلوب منه فور الانتهاء من قراءة المهمة بيسر، كما يجب القيام بتجريب المهمة ميدانياً على مجموعة من الطلاب وذلك لتعرف الأخطاء أو النواقص فيها، وفي حالة أن التجريب غير ممكن أو غير عملي يكتفى باستشارة زميل مهنة لمعرفة رأيه وتقييمه للمهمة، فإذا ما أشار التجريب الميداني أو استشارة زميل المهنة إلى إمكانية التحسين في المهمة، يجب القيام بذلك قبل استخدامها لغرض رصد الدرجات للطلاب.

ج- بناء خطة التصحيح:

عادة ما يستخدم المعلم في اختبارات الورقة والقلم ما يعرف بمفتاح الإجابة، وهو بيان بالإجابة الصحيحة لكل سؤال في الاختبار، ولكن الحال في اختبارات الأداء يختلف نوعا ما، فتحديد إجابة واحدة صحيحة غير ممكن في كثير من الأحوال، وبذلك نستخدم بدلاً من ذلك معايير التصحيح أو خطة التصحيح scoring rubrics، وهي عبارة عن مجموعة من القواعد أو القوانين أو المعايير أو الضوابط مقسمة إلى مستويات حسب درجة دقتها أو صحتها أو قربها من الهدف الذي تقيسه المهمة، بحيث يقابل كل مستوى درجة معينة، وعادة ما تكون تلك المستويات بين 3 و 7 مستويات، وتعمل خطة التصحيح كمرشد للمعلم أثناء التصحيح لضمان استخدام نفس المعايير أو الضوابط في تصحيح كافة الأوراق، وخطة التصحيح من الممكن أن تكون على هيئة قوائم شطب أو مقياس تقدير.

وتقترح بروكهارت (Brookhart,1993) اتباع الخطوات أدناه عند بناء خطة او معايير التصحيح لاختبارات الأداء:

1- حدد معايير الجودة التي يجب توافرها في عمل الطالب المتمكن من المهارة موضع القياس.

2- حدد نوع خطة التصحيح (تفصيلية أم كلية) وكما يلي:

أ- خطة التصحيح التفصيلية تقوم على تفكيك المهارة موضع القياس إلى مهارات صغيرة ومنفصلة بحيث يتم تقييم الطالب في كل مهارة على حدة، فعند تقييمنا لموضوع تعبير مثلاً نقوم بتجزئة مهارة التعبير إلى مهارات منفصلة تتكون من سلامة اللغة، وترابط الأفكار، والتمكن من علامات الترقيم بحيث يحصل الطالب على درجة منفصلة لكل مهارة من هذه المهارات، وتمكن خطة التصحيح التفصيلية المعلم من الحكم على جزئيات أداء المهارة موضع القياس كل على حدة وتخصيص درجة منفصلة لكل جزء.

ب- خطة التصحيح الكلية تدمج المعايير المستخدمة في الحكم على أداء المهارة في مقياس متدرج واحد، بحيث يتم تقييمها ككل، فالنظرة هنا ليست إلى الجزئيات وإنما إلى المهارة موضع القياس ككل متكامل، ففي مثالنا حول مهارات التعبير، يتم دمج سلامة اللغة،

وترابط الأفكار، والتمكن من علامات الترقيم في خطة واحدة، بحيث يتم مراعاة كل تلك المهارات ولكن ليس بشكل منفصل.

3- قم بتحديد معايير أفضل مستوى أداء، أي الشروط التي يجب أن تتوافر في الإجابة التي تحقق الأهداف المقاسة، أو تقترب كثيراً من تحقيقها، ومعايير أسوأ مستوى أداء، أي المستوى الأدنى من الإجابة المقبولة للطالب. أعطي معايير أفضل أداء الدرجة الأعلى في التقدير (5 مثلاً) ومعايير أسوأ مستوى أداء الدرجة الأدنى في التقدير (1 مثلاً)

4- بعد تحديد معايير الأداء المتميز، ومعايير أداء الحد الأدنى المقبول للإجابة، نستطيع أن نحدد معايير الأداء الوسط بين أفضل وأسوأ أداء وذلك بمقارنة معايير هذين المستويين وتحديد معايير الأداء الوسط بينهما، وبذلك نعطي هذا المستوى الدرجة الوسطى وهي (3 مثلا).

5- ولكي نحدد مستوى الأداء الثاني، ما علينا سوى مقارنة مستوى الأداء الأول والذي أعطيناه الدرجة (5) ومستوى الأداء الثاني والذي أعطيناه الدرجة (3)، وبذلك يحصل هذا المستوى على الدرجة (4)، وبنفس الطريقة نقارن مستوى الأداء الثالث ومستوى الأداء الأدنى لتحديد مستوى الأداء الثاني، ويحصل هذا المستوى على الدرجة (2).

6- تتبع الخطوات في 4، 5 أعلاه لإضافة مستويات أداء أكثر إذا لزم الأمر.

وغالباً ما يُكتفى بخمسة مستويات حتى تكون خطة التصحيح معقولة ويسهل التعامل معها أثناء عملية التصحيح، لأنه كلما زادت المستويات، تصبح الخطة مضللة نوعاً ما، ولربما أثر ذلك على ثبات المصححين.

ولكي يستطيع المعلم بناء معايير أداء جيدة، يرى إيريزيان (Airasian, 1997) ضرورة أن يقوم المعلم بتنفيذ المهمة بنفسه، لكي يتحقق من جودة المهمة في قياسها للهدف المحدد، ولكي يحدد معايير الأداء الجيد، كما يمكنه الاستعانة بالطلاب لتحديد جوانب السلوك التي يجب أن تتضمنها إجابة كل مهمة من مهمات الأداء. ويرى كذلك ضرورة كتابة المعايير بلغة سهلة وواضحة لا تقبل التأويل، ومراجعة المعايير بشكل مستمر للتأكد من خلوها من التحيز bias.

وتجب الإشارة هنا إلى أن خطة أو معايير التصحيح يمكن أن تعطى للطالب مع المهمة أثناء تأديته لها، لأنها في الحقيقة تحدد للطالب معايير الأداء الجيد، وبذلك فهي تعينه على إكمال المهمة بشكل مرضي.

تصحيح اختبارات الأداء:

بعد تطبيق مهام الأداء على الطلاب، تأتي عملية التصحيح والتي يتم فيها مقارنة سلوك الطالب في المهمة بالمعايير التي تم تحديدها والاتفاق عليها مسبقاً، ولعل موضوع الثبات يعد من القضايا الشائكة في تصحيح اختبارات الأداء، سواء كان التصحيح من قبل معلم أو مصحح rater واحد، أو من أكثر من مصحح، والثبات بمعناه الواسع هو حصول السلوك موضع القياس على نفس الدرجة تقريباً بغض النظر عن الطالب الذي أدى المهمة، وبغض النظر عن المصحح، بمعنى آخر فالثبات reliability يعني الاتساق consistencyفي الحكم على السلوك أيا كان مصدر ذلك السلوك، أو من يقوم بالحكم عليه أو تقديره، فهو الابتعاد إن لم يكن كلياً فبشكل كبير عن الذاتية في التصحيح.

فعند قيام مصحح واحد بتقدير درجات السلوك المختلفة لجميع الطلاب، تظهر قضية ثبات المصحح مع نفسه intra-rater reliability، والمتعلقة بمدى الاتساق في تقدير الدرجة بين عينات السلوك المتشابهة بغض النظر عن الطالب المؤدي للسلوك أو المهمة، أي أنه لا يوجد تفاوت في تقدير الدرجات لنفس السلوك باختلاف الطلاب، فوجود مثل هذا التباين سوف يؤدي بالتالي إلى تدني معامل الثبات.

وعند قيام أكثر من مصحح بتقدير الدرجات للطلاب، تبقى قضية ثبات المصحح مع نفسه ماثلة، ولكن تظهر قضية أخرى وهي ثبات المصححين in-ter-rater reliability، أو الاتساق في تقدير الدرجات لنفس السلوك بين المصححين، فمن المهم بأن يقوم جميع المصححين بتطبيق خطة أو معايير التصحيح بنفس الدقة، وأي تفاوت في هذا الجانب بين المصححين سوف يؤدي إلى تدني معامل الثبات.

ومن هنا تظهر الحاجة إلى تدريب المصححين على دقة تطبيق خطة أو معايير التصحيح قبل الشروع في عملية التصحيح، ويتطلب ذلك الآتي:

(1) **التهيئة لمهمة التقييم**: وتشمل إعطاء المصححين فكرة عامة عن السلوك المقاس، والمعايير المستخدمة للحكم على السلوك، والغرض أو الأغراض

من التقويم، وأوجه استخدامات الدرجات التي سيحصل عليها الطلاب (الدوسري، 2004).

(2) **توضيح معايير تقدير الدرجات للمصححين**: ويتم فيها مناقشة موسعة لمستويات التقدير المختلفة، وتوضيح الفروق بينها، ويفضل أن يصاحب ذلك تقديم أمثلة أو نماذج من إجابات الطلاب لكل مستوى (الدوسري، 2004).

(3) **تدريب المصححين على تطبيق معايير أو خطة التصحيح**: ويتم ذلك عن طريق قيام كل المصححين بتصحيح مجموعة من الأعمال بشكل منفرد، وبعد ذلك يتم مقارنة الدرجات المعطاة لكل عمل بواسطة مختلف المصححين، فإذا ما وجد اختلاف في تقدير الدرجات لنفس الطالب، يجتمع المصححين لمناقشة مبررات إعطاء الدرجة للوصول إلى اتفاق مرضي بين المصححين.

(4) **التحقق من ثبات المصححين**: بعد الاتفاق المرضي بين المصححين على تطبيق معايير التصحيح الذي تم في الخطوة 3 أعلاه، يقوم كل مصحح بتصحيح مجموعة من الأعمال (بين 10 إلى 15) ورقة بشكل منفرد، ويتم بعد ذلك حساب معامل الارتباط بين درجات المصححين، ويمكن أن يتم ذلك باستخدام معامل ارتباط بيرسون.

(5) **الشروع في عملية التصحيح الفعلية**: بعد التحقق من ثبات المصححين، تبدأ عملية التصحيح الفعلية لأعمال الطلاب، على أن يتخلل ذلك مناقشة الأعمال الحدية Borderline cases، أي الأعمال التي تكون في المنطقة الحرجة بين الرسوب والنجاح، والاتفاق على تصنيفها.

مميزات اختبارات الأداء:

1- تتميز اختبارات الأداء بمستوى صدق عالي من حيث أن ما يتم قياسه هو شبيه أو جزء مما يصادفه الطالب في الحياة خارج المدرسة، وبذلك فهي تجعل الفجوة بين ما يتعلمه الطالب في المدرسة وما هو موجود في بيئة العمل شبه منعدمة.

2- استنتاج مدى معرفة وقدرة الطالب على أداء مهارة معينة يتم عن طريق قياس الإجراءات (أو الطريقة) process أو المنتج product أو كلاهما.

3- تقيس اختبارات الأداء كل من المعلومة والمهارة والقدرة في نفس الوقت، فهي تجمع الجانب النظري والتطبيقي.

4- تنسجم اختبارات الأداء مع نظرية التعلم الحديثة والتي تتمركز حول:

أ- وجوب أن يتم التعلم من خلال ما سبق تعلمه من مهارات ومعلومات.

ب- تأصيل التعلم بحيث يحدث في مواقف حقيقية وشبيهة بالواقع.

ج- قيام الطلاب ببناء معارفهم عن طريق خبرات مباشرة.

5- تعمل اختبارات الأداء على تحفيز الطلبة على التعلم وجعله ذا معنى بالنسبة لهم (Parke & Lane, 2007)

6- أثرت اختبارات الأداء على طرائق التدريس التقليدية، واسهمت في ظهور طرائق تدريس اكثر ابتكاراً .(Lane & Others, 2002; Parke & Others, 2006)

عيوب اختبارات الأداء:

1- صعوبة بناء مهام أو أسئلة أداء جيدة، حيث أنها تستغرق وقت وجهد كبيرين.

2- صعوبة بناء خطة تصحيح تكون على درجة عالية من الثبات.

3- تستغرق وقتاً وجهداً كبيرين من الطالب للإجابة عليها، ومن المعلم لتصحيحها.

4- غالباً ما يكون ثبات المصححين ضعيفاً، ولكن إذا ما وضعت معايير التصحيح بعناية فيمكن الوصول إلى مستوى عال من الثبات.

5- عدم تمثيل محتوى المادة بشكل جيد، فنظراً للجهد والوقت المطلوبين لأداء مهام إختبارات الأداء، يكون عدد المهام في الاختبار الواحد قليلة.

6- اختبارات الأداء من الممكن أن تكون محبطة للطلاب ضعيفي التعلم لما تطلبه من معرفة جيدة بمحتوى المادة، وكذلك القدرة على أداء المهارة.

الملف الوثائقي Portfolio

عندما تذهب لزيارة طبيبك المختص، عادة ما يكون أمامه الملف الخاص بك، والذي يحتوي على تاريخ مفصل لزياراتك السابقة للمستشفى، من حيث الأعراض التي عانيت منها سابقاً، ونوعية الأمراض التي انتابتك، وتقارير الفحوصات التي أجريت لك؛ وبعد أن تسرد للطبيب ما تشتكي منه وسبب زيارتك له هذه المرة، يقوم هو بتصفح الملف، محاولاً الاستفادة منه في تشخيص حالتك الراهنة، فالملف الوثائقي الذي نقصده شبيهاً إلى حد كبير بملف الطبيب وخاصة في الهدف، وإن اختلف عنه في المضمون؛ فالملف الوثائقي عبارة عن تاريخ مفصل للطالب يشتمل على معلومات عنه، وعينات من أعماله، وتقارير كتبها الطالب عن نفسه، وتقارير أخرى كتبها المعلم عنه، والهدف منه مساعدة المعلم على تقويم تحصيل الطالب وتشخيص نواحي القوة والضعف لديه.

ظهر الملف الوثائقي أول ما ظهر في الصناعة وشاع تداوله فيها، حيث كان يحتفظ العاملون في المصنع بعينات من الأعمال التي يقومون بإنتاجها خلال فترة عملهم في المصنع، لإبرازها لأصحاب الأعمال عندما يقومون بتصنيفهم وتقويم أدائهم؛ حيث تبرز تلك العينات سلسلة المهام والأعمال التي يستطيع أن يقوم بها العامل وجهوده في الرقي بأدائه، وتقدمه في العمل، والتي عادة ما يتم وضعها في ملف أو حقيبة.

لقد حقق الملف الوثائقي نجاحاً ملحوظاً في الصناعة، في حين أنه لم يحقق النجاح نفسه في العملية التعليمية، ويكمن جزء من المشكلة في أن الأهداف التربوية تكون أحياناً غامضة ومفتوحة أمام تأويلات مختلفة، بينما تكون المعايير في المواقع الصناعية المراد قياسها أكثر وضوحاً وأكثر تحديداً، كما أن المقاييس المستخدمة في الصناعة غالباً ما تكون واضحة وعلى درجة عالية من الصدق والثبات، في حين أن معظم المقاييس التربوية تعاني من ضعف في ثباتها وصدقها.

تعريف الملف الوثائقي

على الرغم من حداثة استخدامه في التربية إلا أن هناك تعريفات كثيرة

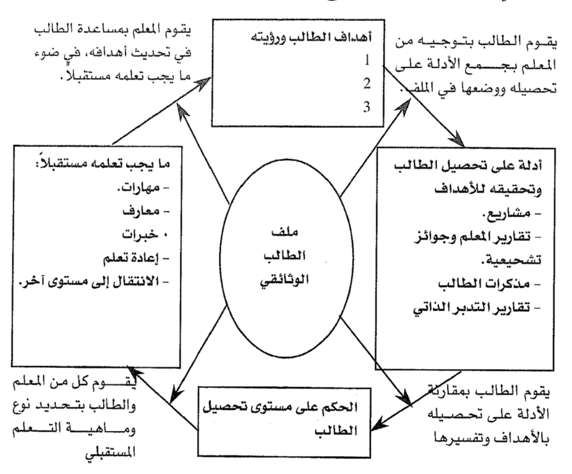
للملف الوثائقي، فرّى شولمان (1992) يعرفه بأنه تاريخ موثق ومنظم لمجموعة من الإنجازات المتدرب على انتقائها والمختارة بعناية، والمدعمة بكتابات تأملية وحوارات هادفة. ويعرفه المختبر التربوي لمنطقة الشمال الغربي الأمريكية (NREL, 1991) بأنه مجموعة من أعمال الطالب تخدم غرضاً معيناً تبين مدى تقدم الطالب أو تحسنه في مجال أو مجالات معينة. ويرى إريزيان (Airasian, 1997) بأنه مجموعة من إنتاجات الطالب أو أعماله. أما نيتكو (Nitko, 2004) فيعّرفه بأنه مجموعة محددة من أعمال الطالب تستخدم بغرض تقديم أفضل الأعمال للطالب، أو توضيح نموه الأكاديمي في فترة زمنية معينة. وتُعرفه عيد (2006) بأنه حصيلة ذات أهداف وأغراض عن اعمال الطالب تمثل جهوده التي قام بها ونشاطاته التي انجزها، كما توضح مدى ما احرزه من تقدم أو نمو دراسي، وما حققه من انجازات في هذا الاتجاه، وذلك في مجال ما من المجالات الدراسية، أو موضوع ما من موضوعات الدراسة.

على أساس ذلك يمكن القول بأن الملف الوثائقي هو عملية تجميع مستمرة لعينات من الأعمال التي ينجزها الطالب بحيث يضم مجموعة من الأعمال التي تعكس تقدمه الأكاديمي خلال فترة زمنية محددة، ومجموعة أخرى تمثل أفضل إنتاجاته، على أن يصاحب ذلك تقارير تقويم وتدبرٍ ذاتي يقوم بكتابتها الطالب في فترات زمنية محددة.

إذا ما هو الجديد في الملف الوثائقي؟ فلربما تساءل البعض بأن وضع أعمال الطالب في ملف أو حقيبة ليس بالشيء الجديد، فالطالب يعمل ذلك منذ قديم الزمان، ولكن الجديد في ذلك يكمن في أمرين اثنين، أولهما المشاركة الفاعلة للطالب في عمل الملف وإدارته وتنظيمه، فالطالب هنا يتحمل المسؤولية التامة - بتوجيه وإرشاد من المعلم - عن ملفه ومحتوياته، فهو من يقوم باختيار محتويات الملف وفق معايير محددة مسبقاً، ومتفق عليها مع المعلم، والطالب هو المسؤول عن تنظيم ملفه وتبويبه؛ بمعنى آخر يتحمل الطالب كامل مسؤولية ملفه الوثائقي، ويقتصر دور المعلم على توجيه الطالب في اختيار الأعمال التي يمكن أن يتضمنها الملف. أما الأمر الثاني والأهم فهو متعلق بقيام الطالب بالتفكير التدبري reflective thinking في تعلمه، أو ما

يعرف بما وراء المعرفة metacognition، حيث يلزم الطالب بكتابة تقارير فترية حول موضوعات الملف، يقيّم فيها تقدمه نحو الأهداف، ويبن فيها نقاط القوة ونقاط الضعف لديه، ويرسم خطط تعلمه المستقبلي التي تعزز نقاط القوة وتعالج نقاط الضعف لديه؛ فالملف الوثائقي ليس وعاء لتخزين أعمال الطالب، بل هو أداة تعلم فاعلة.

ويمكن النظر إلى الملف الوثائقي على أنه نوع من التقويم المستمر للطالب، فهو عملية مستمرة يتم فيها تحديد الأهداف، وتقويم مدى تحققها، وتحديد ما يجب تعلمه مستقبلاً؛ كل ذلك يتم بمشاركة فاعلة من قبل الطالب، ويمكن اعتبار الملف الوثائقي أداة تقويم مستمر، فإذا ما أخذنا نموذج هارلن (Har-len,2000) للتقويم المستمر، نجده ينطبق تماماً على خطوات تنفيذ الملف الوثائقي، والشكل (8-2) يوضح ذلك.



الشكل (8-2) : دور الملف الوثائقي كأداة تقويم مستمر طبقاً لنموذج هارلن

خطوات إعداد الملف الوثائقي:

يتطلب الملف الوثائقي جهداً كبيراً لإعداده وإدارته، وتجربة الملف الوثائقي لن تكلل بالنجاح إذا ما أراد المعلم أن يتحمل مسؤوليتها بالكامل، لأنه لا يخفى على أحد حجم العمل المطلوب لإعداد ملف واحد، فكيف بالأمر في إعداد عشرات الملفات؛ ومن هنا يجب على كل طالب أن يتحمل كامل مسؤولية ملفه الوثائقي، بدءاً من انتقاء مقتنياته، وانتهاءً بإدارته وحفظه، على أن يتم ذلك تحت توجيه وإشراف من قبل المعلم، وفيما يلي نستعرض خطوات إعداد الملف الوثائقي التي قد تعين الطالب والمعلم في عملية إعداد الملف.

1- تحديد الهدف من الملف:

إن تحديد الهدف من الشيء، يعني رسم ملامح العمل لذلك الشيء، وبذلك فتحديد أهداف التعليم وغايات المنهج الدراسي التي يخدمها الملف، سوف تؤدي بالتالي إلى تحديد ماهية الأعمال التي يجب إدراجها في الملف، فالملف الذي يجب أن يظهر نمو وتطور التعلم لدى الطالب يختلف في محتوياته عن الملف الذي يركز على عرض أفضل ما توصل إليه الطالب. وتحديد الهدف من الملف يجب أن يشتمل على تحديد نوعية وآلية التقويم (تكويني أم ختامي) الذي سوف يستخدم للحكم على الملف، كما يجب تحديد المستهدف من الملف، فهل سيقتصر تداول الملف على الطلاب والمعلمين، أم أنه سوف يعرض على أولياء الأمور أيضاً.

2- تحديد محتويات الملف:

إن تحديد أنواع المعرفة والمهارات والقدرات التي ستشكل محتويات الملف وتبويب وتنظيم الملف من الضروري بمكان للطالب؛ فنوعية الأعمال التي تعكس تطور أداء الطالب في المادة الدراسية، يختلف عن تلك الأعمال التي تعكس ما توصل إليه الطالب من إتقان للمادة العلمية. وبالرغم من أن الأعمال التي يحتويها الملف سوف تختلف باختلاف الهدف منه، إلا أن هناك أساسيات معينة يجب أن تكون جزءاً رئيسياً من كل ملف، نذكر منها ما يلي:

أ- البيانات الأولية للطالب: وهي استمارة موحدة لجميع الطلاب، وتمثل الصفحة الأولى في الملف، وتشمل هذه البيانات: الإسم، وتاريخ الميلاد، وعنوان الطالب، ووسائط الاتصال المختلفة (تلفون المنزل، تلفون الأم والأب، البريد الالكتروني)، وفصيلة دم الطالب، وطوله ووزنه وغيرها من البيانات التي قد يرى المعلم أنها ضرورية.

ب- أهداف الطالب ورؤيته: ويشتمل ذلك صياغة أهداف الطالب التعليمية، ورؤيته المستقبلية، كأن يذكر المهارات التي يحاول اكتسابها في المادة مثلاً، وماذا يريد أن يكون مستقبلاً، فليس بالضرورة أن تكون تلك الأهداف طموحة أو معقدة، فليس الهدف منها تحديد مستقبل الطالب بقدر ما هو تدريب الطالب على صياغة أهداف معينة والعمل على تحقيقها، وبذلك يمكن أن تكون تلك الأهداف بسيطة ومتواضعة.

ج- عينات أعمال الطالب: في ضوء الأهداف التي حددها الطالب، ووفقاً لمعايير معينة يتم الإتفاق عليها بين المعلم والطالب، يقوم الطالب بوضع عينات من عمله في الملف. فإذا كان الهدف من الملف إبراز ما وصل إليه الطالب من مستوى في المادة فيجب أن يشتمل الملف على عينات من أفضل انتاجات الطالب، وإذا كان الهدف من الملف أن يبرز الطالب مستوى تقدمه في المادة، فهنا يجب أن يشتمل الملف على عينات (مسودات) أولية من أعمال الطالب، وأعمال شبه منتهية، وأخرى في صورتها النهائية. ويجب الإشارة هنا إلى ضرورة أن يتضمن الملف عينات من الأعمال المختلفة التي قام بها الطالب، وألا يقتصر ذلك على جانب معين من المادة.

3- تقارير التدبر الذاتي **Self-reflection reports**

من أهم أهداف الملف الوثائقي - إن لم يكن أهمها على الإطلاق - هو دفع الطالب إلى التفكر فيما يتعلمه، وفي الحقيقة هذا ما يميز الملف الوثائقي الذي ننشده عن أي أداة أخرى لتجميع أعمال الطالب. حيث يطلب من الطالب على فترات زمنية محددة بأن يقوم بمراجعة محتوى ملفه في ضوء

الأهداف التي رسمها لنفسه، وتقييم مدى تقدمه نحو الأهداف المرسومة، وما هي نقاط القوة والضعف لديه، وما هي خططه للعمل المستقبلي، أي كيف سيعزز نقاط القوة، ويتلافى نقاط الضعف حتى يصل في النهاية إلى ما يريد، على أن يدون ذلك في تقرير مبسط يُدرج ضمن محتويات الملف.

4- إدارة الملف وتنظيمه:

كما أسلفنا بأن كل طالب مسؤول مسؤولية كاملة عن محتويات ملفه وإدارته، ولكن يبقى على المعلم الإشراف المباشر على هذه العملية، كما يجب عليه أن يخصص الوقت اللازم للطلاب للعمل في ملفاتهم، وما يتضمنه ذلك من انتقاء للأعمال وتحديثها، وتقييم ملفاتهم وكتابة تقاريرهم الذاتية حولها بشكل دوري؛ ويمكن للمعلم أن يخصص حصة واحدة أسبوعياً أو شهرياً (يعتمد على عدد الحصص الأسبوعية للمادة) لهذا الغرض؛ فتخصيص الوقت لا يضمن بأن يعمل الطلاب على تحديث ملفاتهم وتقويمها بشكل دوري فحسب، وإنما يتيح الفرصة للمعلم للاطلاع ومتابعة الطلاب أولاً بأول، كما يستطيع المعلم خلاله من تقييم بعض الملفات وتقديم تغذية راجعة للطلاب، وما يجب الإشارة إليه هنا، هو أن هذا الوقت قد لا يكون كافياً للعمل في الملف بشكل مرضي، وإنما يجب على الطالب تكملته في وقت فراغه.

تقويم الملف الوثائقي:

إن استخدام الملف الوثائقي في التقويم يجب أن يخضع إلى كافة الأسس المعمول بها في الاختبارات، ومن هذه الأسس ضرورة تبليغ الطالب بموعد التقويم حتى يستطيع أن يكون مستعداً له. وتقويم الملف الوثائقي يمكن أن يقوم به المعلم بنفسه، كأن يجمع الملفات من جميع الطلاب في يوم محدد ومعروف لدى جميع الطلاب، ويقوم بعملية التقويم وفقاً لمعايير التقويم المتفق عليها ويمكن أن تكون العملية بمشاركة الطالب، كأن يجلس الطالب مع المعلم أثناء عملية التقويم، ويقدر الدرجة التي يستحقها وفق المعايير المعروضة أمامه، كما يمكن مناقشة الطالب شفهياً عن محتويات الملف.

إن استخدام الملف الوثائقي في عملية التقويم ما يزال بين مؤيد ومعارض، فيرى فريق من التربويين إمكانية تطبيقه في التقويم التربوي لعدة أسباب منها: (Lemahieu, and Others, 1995)

1- يعتبر أفضل من الاختبارات المعيارية والتحريرية، لأنه يقدم معلومات أوفر عن الطريقة التي يفكر بها الطلبة حول المشكلات التي يواجهونها، وطرق حلها.
2- يُمكّن الطلبة من لعب دور أكثر حيوية في عملية التقويم؛ فالطالب يعرف مسبقاً معايير الحكم على الملف، ومعايير اختيار ما يحتويه الملف.
3- يوفر مهام تقويم أكثر شمولية، ويسمح للطلبة بتعلم مواد مفيدة أكثر وذات علاقة أكبر مما هو الحال عند استخدام الاختبارات التحريرية.
4- يتيح الفرصة للمتعلم للتدبر في تقدمه التحصيلي نحو تحقيق أهدافه التي حددها لنفسه.
5- يساعد المؤسسات التربوية في تقويم برامجها وتطوير برامج جديدة قائمة على الكفاءة أو الأداء في مجالات متعددة من خلال ما يقدمه من براهين وأدلة ملموسة عن تعلم الطلاب.

في حين لا يتردد البعض من التصريح عن مخاوفهم حول استخدام الملف في عملية التقويم، ويرجعون تخوفهم هذا إلى قضايا عدة منها:

1- عدم وجود مستوى مقبول من الاتفاق بين المحكمين في وضع الدرجات عند تقويمهم للملف؛ فعادة ما يكون ثبات المصححين متدنياً.
2- صعوبة بناء معايير تصحيح الملفات الوثائقية من قبل المعلم.
3- تستغرق وقتاً كبيراً وجهداً مضاعفاً من قبل المعلم، خاصة وأنها تستخدم بالإضافة إلى وسائل التقويم الأخرى .(Oosterhof, 2001)
4- عدم وجود اتفاق مقبول حول المستويات أو المعايير الواجب استخدامها عند تقويم الملف.
5- الأحكام الصادرة بحق أداء الطالب قد تتغير بوصفها عملاً لعينات من الأعمال المختارة، فإذا ما قام الطالب بعرض عينات أخرى فإن مستواه قد يتغير نزولاً أو صعوداً.

ومن خلال استعراضنا للمخاوف التي أبداها معارضو استخدام الملف الوثائقي لغرض التقويم، يتضح بأن تلك المخاوف في مجملها نابعة من حداثة استخدام الملف الوثائقي كأداة تقويم؛ فالاتفاق بين المصححين يمكن علاجه من خلال تبني خطط أو معايير تصحيح واضحة ومحددة، وان يتم تدريب المعلمين على بناء المعايير، وتدريب المصححين على استخدامها قبل الشروع في عملية التصحيح. أما فيما يتعلق بالاتفاق على المستويات والمعايير الواجب استخدامها عند تقويم الملف، فإن تلك المستويات والمعايير تحددها أهداف الملف والأغراض التي يستخدم من أجلها.

معايير التقويم:

إذا ما نظرنا إلى الملف الوثائقي على أنه اختبار أداء يتطلب من الطالب جهداً كبيراً متواصلاً، وفترة طويلة من الزمن لإنجازه، ويتضمن العديد من المهارات والخبرات، فإن تقويمه لا يختلف كثيراً عن اختبارات الأداء، وبذلك يجب أن يكون لدى المقوم معايير تصحيح متفق عليها لكي تعينه في تقييمه للملفات المختلفة وحتى لا تختلف المعايير عند انتقاله من ملف لآخر، ولقد سبق وأن تناولنا بشيء من التفصيل خطوات بناء المعايير في اختبارات الأداء، حيث يمكن استخدام نفس الخطوات في بناء معايير الملف الوثائقي، على أن تركز تلك المعايير على الغرض الأساسي للملف كونه يعكس مستوى التحصيل الذي وصل إليه الطالب، فهو بذلك يتضمن أفضل أعمال الطالب، أو أنه يعكس التقدم الذي أحرزه الطالب في دراسته للمادة، وهو بذلك يتضمن أعمال في مراحل مختلفة من الإنجاز.

يتعين أن يشمل تقويم الملف الوثائقي المعايير التالية:

1- الإطار التنظيمي للملف: احتواء الملف على العناصر المختلفة (قائمة محتويات، الرؤية والأهداف، عينات من أعمال الطالب، تقارير التدبر الذاتي self-reflection reports، تقارير مكتوبة عن الطالب، شهادات التقدير والجوائز التشجيعية التي حصل عليها الطالب إن وجدت) وحسن

تصنيف وتبويب تلك المحتويات داخل الملف بحيث يسهل تصفحها وتقويم مدى ملاءمتها للأهداف، علماً بأن التركيز هنا على الناحية الشكلية للملف فقط.

2- محتوى الملف: ويجب أن تتناول المعايير في هذا الجانب إيجاد مؤشرات تقييم الجانب النوعي في الملف، بحيث يتم مقارنة محتويات الملف ومدى تحقيقها لأهداف الطالب المختلفة، أو مدى تقدمه نحو تحقيق تلك الأهداف، ويشمل ذلك:

أ- تقييم عينات أعمال الطالب من حيث شمولها للمواد أو العناصر المختلفة للمادة الواحدة، ومدى تنوعها وارتباطها بالأهداف.

ب- تقييم تقارير التدبر الذاتي من حيث موضوعيتها، واشتمال كل تقرير على العناصر الأساسية من وصف الوضع الراهن لمستوى الطالب، وبيان نواحي القوة والضعف، ورسم خطط التعلم المستقبلي.

ج- تقييم تقارير المعلم حول الطالب وأهم ملاحظات المعلم عليه، ومدى استفادة الطالب من تلك الملاحظات.

د- تقييم مؤشرات التميز من حيث وجود شهادات تقدير وجوائز تشجيعية حصل عليها الطالب.

أهمية الملف الوثائقي:

يستخدم الملف الوثائقي في التربية في مختلف مراحل التعليم، ولكنه أكثر شيوعاً في المراحل العليا بدءاً من الصف العاشر، كما أنه يستخدم في مختلف المواد الدراسية ولأغراض مختلفة كأداة تدريس، وتقويم، وتشخيص (Lock-ledge,1997; Phelps, & others, 1997; Krueger & Wallace, 1996; Jones, 1994, Wolf, 1989) وقد أثبتت الدراسات فاعليته في مختلف جوانب العملية التعليمية التعلمية، وعلى وجه التحديد ادناه اهم استخدامات الملف الوثائقي:

1- استخدام الملف الوثائقي كأداة تخطيط العملية التعليمية- التعلمية، بحيث يكون الملف الوثائقي موجهاً للخبرات التعليمية داخل غرفة الصف،

وهناك دراسات أثبتت فاعلية هذا التوجه في المرحلة الثانوية في مواد الدراسات الاجتماعية (Lockledge,1997)، والكيمياء (Phelps et. al., 1997)، واللغات الأجنبية .(Hayes, 1997).

2- استخدام الملف الوثائقي كوسيلة لجعل الطلاب يتحملون مسؤولية تعلمهم، وإشراكهم في عملية التقويم بفاعلية .(Phelps et. al., 1997).

3- استخدام الملف الوثائقي كوسيلة لتخطيط احتياجات الطلاب، مما ينتج عنه تحسين عملية تخطيط الخبرات التدريسية وتنفيذها .(Hayes, 1997; Jones, 1994).

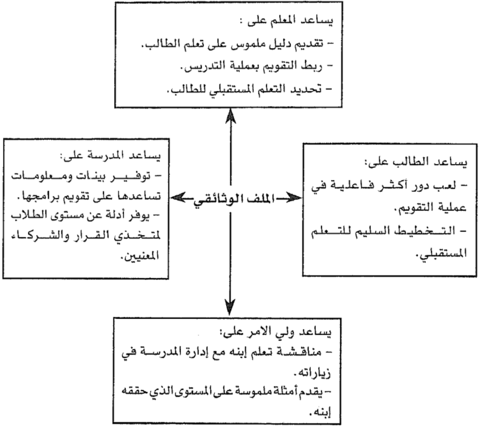
وللملف الوثائقي أهمية كبيرة لأطراف العملية التعليمية التعلمية: الطالب، المعلم، المدرسة، وولي الأمر. فبالنسبة للطالب، يعمل الملف الوثائقي على أن يكون للطالب دور أكثر حيوية في عملية التقويم، حيث أن الأدوات التقليدية تغيب دور الطالب في عملية التقويم، ويقتصر دوره فقط في الجلوس لتأدية الإختبارات، فهو لا يشارك في إعدادها، وتصحيحها؛ كما أنه يشجعه على التفكير فيما يتعلمه من خلال كتابة تقارير التدبر الذاتي الدورية التي تساعده على تلمس نقاط القوة والضعف لديه؛ وأخيراً وليس بآخر فهو يعتبر نقطة انطلاق نحو مواصلة التعلم مستقبلاً، ففي ضوء تقييم الطالب لنقاط القوة والضعف لديه يقوم بتحديد ما يجب تعلمه مستقبلاً، وآلية تعلمه.

أما بالنسبة للمعلم فيقدم الملف الوثائقي دليلاً على تعلم الطالب من خلال ما يحويه من مقتنيات متعددة، ويعمل كأساس لمناقشة التقدم الذي يحرزه الطالب في تحصيله الدراسي، مع معلمة أو ولي أمره واستكشاف الأفكار والآليات التي تحسن من أدائه، ويساعد المعلم على ربط عملية التدريس بعملية التقويم والذي قلّ ما توفره أدوات التقويم الأخرى، كما يوفر الملف الوثائقي معلومات للمعلمين الذي سيقومون بتدريس الطالب في الصف اللاحق تمكنهم من معرفة جوانب القوة والضعف لديه.

وبالنسبة للمدرسة كمؤسسة تعليمية يعمل الملف على توفير بيانات ومعلومات تساعد المدرسة على تقويم برامجها وتطويرها، ويوفر أمثلة موثوق

بها عن نوعية الأعمال التي ينجزها الطلاب ومستواها لمتخذي القرار decision makers والشركاء المعنيين stakeholders بالعملية التعليمية.

ويعمل الملف الوثائقي كأساس لمؤتمرات الطالب وولي الأمر student-parent conferencing؛ حيث يقوم الطالب باستعراض ما أنجزه وما يستطيع القيام به لولي أمره خلال زيارته للمدرسة بحضور المعلم، وهو بذلك يعتبر دليلاً ملموساً وعملياً لولي الأمر على التقدم الذي حققه الطالب في المواد الدراسية المختلفة.



الشكل (8-3) استفادة أقطاب العملية التعليمية التعلمية من الملف الوثائقي

مجالات أخرى لاستخدام الملف الوثائقي في التربية

إن استخدام الملف الوثائقي مع الطلاب داخل غرفة الصف ما هو إلا أحد الاستخدامات العديدة التي يمكن استخدام الملف الوثائقي فيها في التربية، وفيما يلي شرح موجز لبعض هذه الاستخدامات:

ملف المعلم: يمكن أن يكون للمعلم ملفه الوثائقي والذي من خلاله يقوم بتوثيق أعماله وانجازاته التي قام بها خلال السنة الدراسية، والغرض منه ذو شقين: الأول يُعنى الاحتفاظ بتلك الأعمال التي بذل فيها المعلم جهداً كبيراً، واستثمر فيها وقتاً طويلا ليتمكن من استخدامها مستقبلاً، فليس معقولا بأن يبدأ المعلم من نقطة الصفر كل مرة يحاول فيها تحضير دروس لطلابه، أو كتابة امتحان، أو التخطيط لمشروعات مدرسية دون الاستفادة من تجارب الماضي. أما الشق الثاني فيعنى باستخدام الملف لغرض التقويم، كأن يعرض المعلم ملفه للمرشد أو الموجه التربوي أو مدير المدرسة ليتم تقييمه في ضوئه.

ويمكن أن يشتمل ملف المعلم الوثائقي على:

(1) السيرة الذاتية للمعلم.
(2) نموذج من تحضير الدروس.
(3) تقارير الزيارات الصفية لمدير المدرسة.
(4) تقارير المرشد أو الموجه التربوي.
(5) وصفاً مختصراً أو خططا لمشروعات قام بتنفيذها أو الإشراف عليها.
(6) تجارب عملية قام بإجرائها أو وصفاً لها.
(7) نماذج اختبارات ورقة وقلم أو تقويم بديل قام بإعدادها.
(8) برمجيات حاسوبية تعليمية قام بتصميمها.
(9) دراسات أو بحوث قام بها أو شارك فيها.
(10) شهادات تقدير أو جوائز حاز عليها على مستوى المنطقة التعليمية أو الوطني.

ملف مدير المدرسة : يكمن الهدف في قيام مدير المدرسة من إعداد ملف وثائقي في إبراز ما قامت به المدرسة خلال السنة الدراسية بشكل عام، ليتم تقديمه للموجه الإداري، أو للجنة تقويم الأداء المدرسي التي تزور المدرسة عادة في نهاية العام الدراسي، كما يمكن الاستفادة من مقتنيات الملف في التخطيط والإعداد لمشروعات الأعوام القادمة، خاصة وأن المدير سوف يحصل على ملاحظات الموجه الإداري ولجنة التقويم حول موضوعات الملف، الأمر

الذي سوف يسهم في تطوير الأداء مستقبلا. ويمكن أن يشتمل الملف الوثائقي لمدير المدرسة على:

(1) بيانات عن المدرسة (أعداد الطلاب، والشعب الدراسية، وبيان بأسماء المعلمين وأعدادهم حسب التخصص، والعاملين بالمدرسة، ومرافق المدرسة المختلفة، وخارطة بالمناطق التي تخدمها المدرسة).

(2) المشروعات التطويرية التي قامت بها المدرسة سواءً فيما يتعلق بتطوير العملية التعليمية التعلمية كإقامة دورات الإنماء المهني للمعلمين أو مشروعات خدمة المجتمع المحلي، على أن يشتمل ذلك على شرح مختصر لكل مشروع مع تقديم مؤشرات ملموسة على تنفيذه.

(3) تقارير التقويم الذاتي الدوري للمدرسة.

(4) التقارير الدورية للموجه الإداري.

(5) تقارير لجنة تقويم الأداء المدرسي.

(6) محاضر اجتماعات مجلس المدرسة ومجلس المعلمين وأولياء الأمور.

(7) الجوائز التشجيعية وشهادات التقدير التي حصلت عليها المدرسة.

ملفات الطلاب المعلمين: يعمد طلاب كليات التربية أثناء دراستهم في كليات التربية لإعداد ملفات وثائقية يجمعون فيها ما أنجزوه خلال سنوات دراستهم بهدف إبراز أنفسهم بشكل أفضل خلال مقابلات التعيين التي عادة ما تتم فور تخرجهم، ويتضمن الملف:

(1) السيرة الذاتية للطالب.

(2) إفادة التخرج من المؤسسة التي درس فيها.

(3) كشف بالمواد التي درسها والتقدير الذي حصل عليه في كل مادة معتمد من الجهة المانحة.

(4) نماذج لدروس قام بإعدادها أثناء دراسته للمقررات المختلفة أو التربية العملية.

(5) نماذج لبرمجيات حاسوبية تعليمية قام بتصميمها.

(6) دراسات وبحوث قام بها أثناء دراسته.
(7) استمارات تقييمه في دروس التربية العملية.
(8) شهادات تقدير أو جوائز تشجيعية حصل عليها.

معوقات استخدام الملف الوثائقي:

لا أحد ينكر حجم العمل المطلوب من الطالب وربما المعلم في الملف الوثائقي، وقد أشارت هايز (Hayes, 1997) إلى عدد من المعوقات التي تقف حجر عثرة في استخدامه على نطاق واسع:

1- حجم الأعمال الورقية المطلوب من الطالب القيام بها والمتمثلة في ملء البيانات الأولية، وكتابة تقارير التدبر الذاتي الدورية، وتجميع الوثائق الأخرى التي تدّعم تحقيق الطالب للأهداف.
2- الوقت المتاح للطالب لإنجاز أعمال الملف غالبا لا يتناسب والعمل المطلوب إنجازه، مما يشعره ببعض الإحباط متمثلا في عدم قدرته على إنجاز ما هو مطلوب منه.
3- عدم قدرة الطالب على تقويم مدى تحقق بعض الأهداف، فإما أن تكون تلك الأهداف من النوع الذي يأخذ فترة طويلة جدا لتحقيقها، أو أنها تحتاج لبراهين كثيرة لإثبات تحققها.
4- عدم قدرة الطالب على التدبر والتفكير فيما تعلمه عند كتابته لتقرير التدبر الذاتي، وهذا ما يجعل بعض المعلمين يقاومون استخدام الملف الوثائقي (Jones, 1994).
5- يمكن أن تحتوي تلك الملفات على أعمال منقولة حرفيا من كتب، وخالية من إبداعات الطالب (Gresham, 1992).

مراجع الفصل الثامن

1- الدوسري، راشد حماد (2004). القياس والتقويم التربوي الحديث. عمان: دار الفكر.

2- السعدي، جميل (2002). تقويم امتحانات شهادة الثانوية العامة لمادة التاريخ بسلطنة عمان في ضوء المستويات المعرفية. رسالة ماجستير غير منشورة، كلية التربية، جامعة السلطان قابوس، سلطنة عمان.

3- عيد، غادة خالد (2006). القياس والتقويم التربوي مع تطبيعات برنامج **SPSS**. الكويت: مكتبة الفلاح للنشر والتوزيع.

4- كاظم، علي مهدي (2001). القياس والتقويم في التعلم والتعليم. اربد: دار الكندي.

5- Airasian, P.W. (1997). **Classroom Assessment**. New York: McGraw Hill.

6- Alsarimi, A. M. (2000). **Classroom assessment and grading practices in the Sultanate of Oman**. DIA. A doctoral dissertation, University of Pittsburgh.

7- Barootchi, N., & Keshavarz, M. (2002). Assessment of achievement through Protfolios and teacher - tests. **Educatinoal Research**, 44(3), 279-388.

8- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. **Journal of Educational Measurement**, 30, 123-142.

9- Brookhart, S. M. (2001). **Performance assessment rubrics. Practical Assessment, Research and Evaluation**. Online Journal: http://www.edresearch.org/pare/

10- Brualdi, A. (1998). implementing performance assessment in the classroom. Practical Assessment, **Research and Evaluation**, 6(2).

11- Cizek, G.J. (2001). **Setting Performance Standards: Concepts, Methods, and Persepectives**. Mehwah, NJ: Lawrence Erlbaum Associates Publishers.

12- Gresham, Oregon School District (1992). **Self and peer evaluation**. Available at http://www.ericae.net/file://A/Assessment.

13- Hayes, J. R. (1996). **A new model of cognition and affect in writing**. In C. M. Levy & S. Ransdell (eds)., The Science of Writing. Mahwah, NJ: Erlbaum.

14- Harlen, W. (2000). **Teaching, Learning and Assessing Science 5-12**. London, Paul Chapman Publishing Ltd.

15- Jones, K. (1998). **Alternative assessment**. Available online: http:// www.cbe.abca/b868/assess/ALT_ASSE.HTM (retrieved 8 August 2001).

16- Kruger, B., & Wallace, J. (1996). Portfolio Assessment: Possibilities and Points for Practice. **Australian Scinces Teachers Journal**, 42(1), 26-29.

17- Lane, S; Parke, C.; Stone, C. (2002) The Impact of a state Performance - Based Assessment and Accountability Prgram on Mathematics Instruction and Student Learning: Evidence from Survey Date and School Performance. **Educational Assessment**, 8(4) 279-315.

18- Lane, S., Parke, C., & Stone, C. (1999). **MSPAP Impact Study**, Volume I: Mathematics. Unpublished research report.

19- LeMahieu, P., & Others (1995). Portfolios in Large - Scale Assessment: Dificult but Not Impossible. **Educational Measurement: Issues and Practice,** 14(3) 11-16.

20- Lockledge, A. (1997). Portfolio Assessment in Middle - School and High - School Social Studies Classrooms. **Social Studies**, 88(2) 65-69.

21- McMillan, J. H. (1997). **Classroom Assessment: Principles and Practices for Effective Instruction.** Allyn & Bacon, Needham Heights, MA.

22- Nitko, A (2004). **The Educational Assessment of Students**. University of Arizona, Prentice Hall.

23- Nitko, A & Brookhart, S. (2007) **Educational Assessment of Students.** New Jersy: Pearson Education, Inc., Upper Saddle River.

24- Northwest Regional Educational Laboratory (NREL) Test Center (1991). **Portfolio Resources**. Eric document (ED343910).

25- Oosterhof, A. (2001). Classroom Applications of Educational Measurement. Florida State University, Prentice Hall.

26- Parke, C., & Lane, S. (2007) Sturdents? Perceptions of a Maryland

State Performance Assessment. Elementary School Juurnal, 107(3). 239-269.

27- Parke, C; Lane, S; & Stone, C. (2006). Impact of a state Performance Assessment Program in Reading and writing. Educatinal Research and Evaluation, 12(3). 239-269.

28- Phelpes. A., & Others (1997) Portfolio Assessment in High School Chemistry: One Teacher's Guidelines. **Journal of Chemical Education**, 74(5). 528-31.

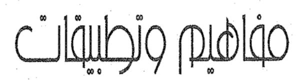
29- Schulman, L. (1991). New Assessment Practices in Mathematics. Journal of Education, 178(1). 61-71.

30- Stiggins, R. J. (1991). **Student-centered classroom assessment**. New York: Merrill.

31- Wolf, D.P. (1989). Portfolio assessment: Sampling student work. **Educational Leadership**, 46(7), 35-39.

الباب الثالث

الإحصاء والبرامج الحاسوبية


مفاهيم وتطبيقات

الفصل التاسع

المفاهيم الإحصائية الأساسية في القياس والتقويم

- أهمية الإحصاء
- تعريف الإحصاء
- علاقة الإحصاء بالقياس والتقويم
- طرائق عرض البيانات
- مقاييس النزعة المركزية
- مقاييس التشتت
- مقاييس العلاقة

9

الفصل التاسع

المفاهيم الإحصائية الأساسية في القياس والتقويم

أهمية الإحصاء

نتيجة للتطورات الحاصلة في جوانب الحياة كافة، نجد الانسان اليوم يتعامل مع الارقام في كل منحى من مناحي حياته؛ فلغة الارقام ومفاهيم الاحصاء مستخدمة في عرض نتائج المباريات الرياضية، وأسواق الأسهم، وأسعار العملات، وكميات الصادرات والواردات، والاحصاءات المستخدمة على نطاق واسع؛ كاعداد العاملين في مؤسسة معينة موزعين على حسب متغير النوع (ذكور- اناث)، والحالة الاجتماعية (متزوج- اعزب)، وحتى الكتب المؤلفة حديثاً تتميز عن الكتب القديمة في إحتواؤها على الجداول والمنحنيات والرسوم البيانية والأرقام. . . . . .الخ.

لقد اصبحت لغة الارقام اليوم، لغة العلوم كافة، وأصبح لزاماً على الفرد معرفة تلك الأرقام وفهمها، من خلال دراسته طرائقاً معينة تعينه على ذلك، وهذه الطرائق تعتمد الاحصاء اساساً لها.

فالاحصاء إذن؛ يحتاج اليه الانسان العادي، والانسان المتخصص في مجال معرفي معين، وإن عدم قدرته على استيعاب البيانات المشار اليها بمفاهيم احصائية معينة، تجعله عاجزاً عن استثمار تلك البيانات في المجالات العلمية والعملية في مختلف جوانب الحياة.

على هذا، فالاحصاء تعتمد عليه مختلف ميادين العلوم؛ إذ يعد وسيلة مهمة في وصف ظواهرها ومفاهيمها والوصول الى نتائج دقيقة يعتمد عليها في اتخاذ قرارات مختلفة في درجة أهميتها.

تعريف الاحصاء

يُعّرف الاحصاء Statistics بأنه عملية جمع البيانات وتبويبها وعرضها وتحليلها واستخراج النتائج والاستدلالات منها لغرض اتخاذ القرارات. يتضمن هذا التعريف الاحصاء الوصفي، والاحصاء الاستدلالي.

فالاحصاء الوصفي Descriptive Statixtics يهتم بجمع البيانات وتبويبها وعرضها في صورة جداول احصائية أو رسوم بيانية.

واما الاحصاء الاستدلالي Infrentian Statistics فيهتم باشتقاق النتائج من دراسة البيانات المتوافرة عن ظاهرة معينة، وفحصها باستخدام وسائل احصائية مبنية على نظرية الاحتمالات الرياضية.

علاقة الاحصاء بالقياس والتقويم

هناك علاقة قوية بين الاحصاء والقياس والتقويم، تتلخص طبيعة هذه العلاقة في كون:

1 - الاحصاء وسيلة لجمع البيانات التي نحصل عليها من عملية القياس عند استخدامنا للاختبارات والمقاييس، وبعد جمع البيانات، يتم تبويبها وعرضها وتحليلها واستنباط النتائج واتخاذ القرارات بناءً على ذلك التحليل.

2 - الاحصاء وسيلة يعتمد عليها المعلم اعتماداً كبيراً، بدءاً باعداده للاختبار، وانتهاءً بتوزيع النتائج على الطلبة.

أولاً: طرائق عرض البيانات Methods of Data Presetation

تواجهنا في الحياة العملية كميات كبيرة من البيانات Data، منها ما يتعلق بأعداد الطلبة في الصف أو المدرسة، او نتائج الامتحانات، او اعداد المرضى والأسوياء، او اعداد المتفوقين والمتأخرين دراسياً وغير ذلك.

ان عرضنا لهذه البيانات بطريقة المقال ضمن التقارير، او الصحف اليومية تكون مملة، ويصعب استيعابها وفهمها، واجراء المقارنات فيما بينها؛ لذا لا بد من عرضها بطريقة مبسطة وشيقة تتيح امكانية التعامل معها بسهولة.

هناك طريقتان لعرض البيانات، أحداهما تفيد مع البيانات الأولية، وثانيهما تفيد مع التوزيعات التكرارية، ولكل منهما طرائق عرض خاصة بها. وفيما يلي تفصيل لذلك.

البيانات الاولية وطرائق عرضها

عندما تكون كمية البيانات صغيرة (أي لا تتجاوز 15 فئة)، أو عندما نريد عرض البيانات كما هي، دون إجراء أي عمليات في عرضها، تسمى بالبيانات الاولية Ungrouped Data.

ان الهدف من عرض البيانات الاولية هو ابراز ملامحها الاساسية بوضوح ودقة، كي يسهل فهمها والتعامل معها.

وهناك طرائق عديدة لعرض البيانات الاولية، يتوقف استخدامنا لطريقة ما دون سواها على نوعية البيانات، والهدف من العرض. ومن هذه الطرائق الجداول، والاعمدة البيانية، والخط المنكسر، والدائرة البيانية. وفيما يلي عرض لكل طريقة.

أ - طريقة الجداول **Tabels Method**

وهي عملية وضع البيانات في جداول دون اجراء أي عمليات عليها، سوى ترتيبها زمنياً، او حسب القيمة، أو أي معيار مناسب لذلك الترتيب.

وكثيراً ما تستعمل الجداول في عرض تغير ظاهرة مع الزمن، أو مع مسميات كالبلدان والمدارس مثلاً، أو مع الزمن والمسميات معاً.

مثال (1)

بلغ عدد طلبة مدرسة معينة في العام الدراسي 2006-2007 (820) طالباً، عدد طلاب الصف الاول 175، والصف الثاني 160، والصف الثالث 155، والصف الرابع 130، والصف الخامس 115، والصف السادس 85.

المطلوب: عرض البيانات السابقة في جدول على حسب تغيرها مع المسميات والظاهرة معاً.

الحل: المسميات هي الصفوف، والظاهرة هي اعداد الطلاب.

الجدول (9-1)

أعداد طلاب مدرسة........للعام الدراسي **2006-2007**

موزعين على حسب الصفوف

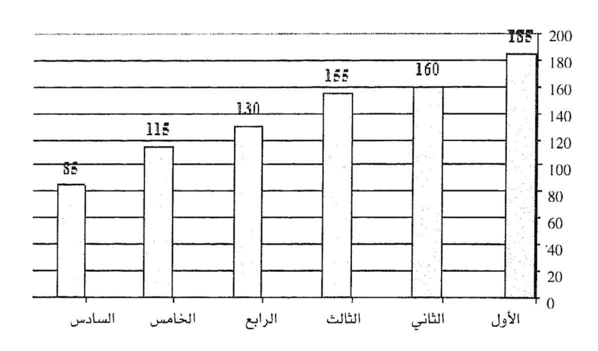
الصف	عدد الطلاب
الأول	175
الثاني	160
الثالث	155
الرابع	130
الخامس	115
السادس	85
المجموع	820

ب- طريقة الاعمدة البيانية **Column Charts Method of Bar Graph**

تتلخص هذه الطريقة بوضع المسميات في محور أفقي أو عمودي، ورسم عمود أو مستطيل على كل مسمى، يكون ارتفاعه متناسباً والقيمة المقابلة لذلك المسمى؛ أي أن جميع الاعمدة لها قاعدة متساوية، واطوالها مختلفة بحسب القيم المناظرة لها.

تهدف هذه الطريقة الى مساعدة القارئ على معرفة الخصائص الجوهرية والمهمة للبيانات بمجرد النظر اليها.

وتستعمل الاعمدة البيانية لعرض تغير ظاهرة مع الزمن، او مع مسميات معينة، أو كليهما معاً؛ حيث يمكن استعمالها للمقارنة بين قيم الظواهر حسب المسميات على مدى عدة سنوات، كأن نقارن بين اعداد الطلبة بحسب تخصصاتهم في الجامعة على مدى خمس سنوات، أو تعرف اعداد الطلبة في مدرسة ما حسب الصفوف. والشكل البياني (9-1) يتضمن اعمدة بيانية لبيانات المثال (1).

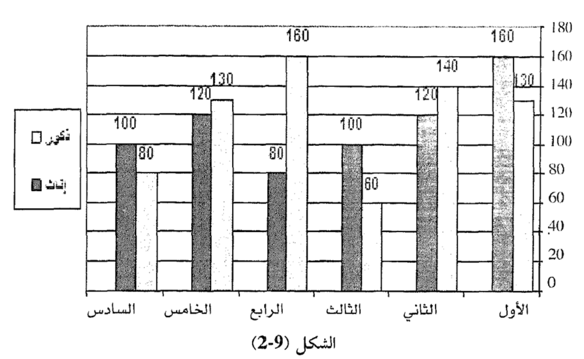


الشكل (9-1)

أعمدة بيانية لأعداد طلبة مدرسة مفترضة للعام الدراسي 2006-2007

ويمكن عرض الأعمدة بشكل متجاور (متلاصق)، مما يسمح بالمقارنة المباشرة بنحو افضل وأدق مما لو كانت الاعمدة متباعدة. على أن يبدأ رسم العمود في كلا الحالتين من الصفر، كي لا تعطي الاطوال النسبية للاعمدة في حالة البداية بغير الصفر انطباعاً خاطئاً عند مقارنتها ببعضها.

كما يمكن كتابة الرقم المعبر عن حجم العمود اكثر من متغير في رسم بياني واحد، كأن نعرض أعداد الذكور والاناث، ونميز أعمدة الذكور بلون، أو شكل مختلف عن اعمدة الاناث. وتسمى الاعمدة في هذه الحالة بالاعمدة المزدوجة. والشكل (9-2) يبين اعمدة بيانية لاعداد الذكور والاناث.



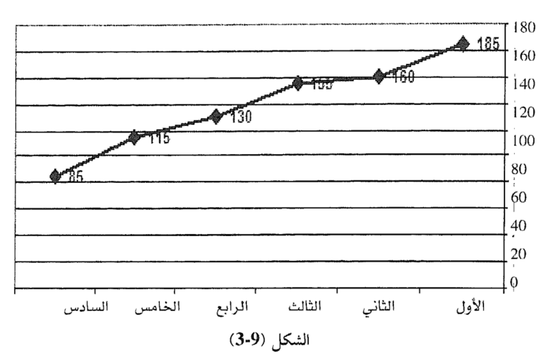
الشكل (9-2)

أعمدة بيانية مزدوجة لأعداد الذكور والإناث

وفي بعض الحالات يمكن عرض أكثر من متغيرين في رسم بياني واحد، كأن نعرض نتائج الامتحانات للناجحين والراسبين والمكملين في كل صف دراسي، أو ما شابه ذلك.

ج- طريقة الخط المنكسر أو المنحني

تستعمل هذه الطريقة لعرض البيانات الناتجة عن تغير ظاهرة، أو عدة ظواهر مع مسميات، أو مع الزمن، أو كليهما معاً، مثل تغير درجة حرارة مريض مع الساعات، أو تغير اعداد الطلبة مع السنوات، أو تغير اعداد الطلبة حسب الكليات على مدى فترة زمنية محددة. وفي حالة المثال (1) يمكن عرض البيانات بطريقة الخط المنكسر كما في الشكل (9-3).



الشكل (3-9)

الخط المنكسر لأعداد طلبة مدرسة مفترضة حسب الصف

ويمكن عـرض عـدة ظواهر في رسم بيـاني واحـد، على أن يتم تمييـز الخطوط بألوان أو اشكال مختلفة.

وأما الخط المنحني فيتم الحصول عليه من خلال تمهيد الخط المنكسر باليد، حتى يصبح منحنياً.

د- طريقة الدائرة البيانية **Pie Charts Methods**

تستخدم طريقة الدائرة البيانية، او المخططات الدائرية لغرض المقارنة بين كميات مختلفة لمتغير واحد، كأن نعرض اعداد طلبة الجامعة حسب الكليات، او اعداد طلبة مدرسة معينة حسب الصفوف. ويتم ذلك بتقسيم الدائرة الى أجزاء تتناسب مساحاتها وعدد الطلبة في كل كلية أو كل صف دراسي.

يتم تقسيم الدائرة على اساس ضرب زاوية مركز الدائرة وهي °360 في نسبة كل جزء للمجموع الكلي. وبالنسبة للمثال (1) يمكن عرضه بطريقة الدائرة البيانية على وفق الخطوات التالية:

$$\text{قياس الزاوية} = \frac{\text{عدد طلبة الصف}}{\text{المجموع الكلي}} \times 360$$

1 - زاوية الصف الأول= $\frac{175}{820} \times 360 = 76.83 = 77$

2 - زاوية الصف الثاني= $\frac{160}{820} \times 360 = 70.24 = 70$

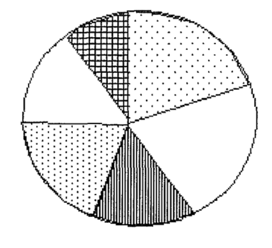
3 - زاوية الصف الثالث= $\frac{155}{820} \times 360 = 68.05 = 68$

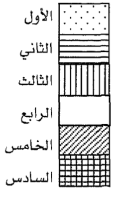
4 - زاوية الصف الرابع= $\frac{130}{820} \times 360 = 57.07 = 57$

5 - زاوية الصف الخامس= $\frac{118}{820} \times 360 = 50.49 = 51$

6 - زاوية الصف السادس= $\frac{85}{820} \times 360 = 37.32 = 37$

المجموع $360°$





الشكل **(4-9)** الدائرة البيانية لاعداد طلبة مدرسة............

التوزيعات التكرارية وطرائق عرضها

عندما يكون عدد البيانات كبيراً (اكثر من 15 قيمة او عدد)، أو عندما يريد الباحث اختزال البيانات الاولية لهدف ما، توضع في فئات Intervals، وكل فئة تضم مجموعة متقاربة من القيم، وتقابلها تكراراتها. إن هذا الاجراء يعرف بالتوزيع التكراري Frequency Distribution

والتوزيع التكراري يكشف عن بعض الصفات البارزة، أو الخصائص العامة التي لا تظهر في البيانات قبل تنظيمها، وفي التوزيع التكراري نلخص البيانات ونعرضها بطريقة مبسطة تسهل فهمها؛ بحيث لا تخسر البيانات من اهميتها إلا الشئ اليسير، وربما لا تخسر شيئا.

ويتعين أن يراعى في فئات التوزيع التكراري ألا يكون عددها كبيراً، فتنتهي الحكمة من عملية التجميع، وألا يكون عددها صغيرا، فتضيع معالم التوزيع، ويفقد الكثير من تفاصيله، ويفضل أن يتراوح عدد الفئات في التوزيع بين 5 الى 15 فئة.

مثال (2):

اختبر معلم طلابه في مادة التاريخ وعددهم (25) طالباً، وكانت درجاتهم على النحو التالي:

10	9	5	6	8
4	3	0	1	2
8	9	8	8	7
6	3	7	4	6
5	6	7	9	8

المطلوب: عرض الدرجات السابقة في جدول تكراري.

الحل:

نقوم اولاً بتحديد اصغر قيمة واكبرها وهي: صفر و 10 .

بعدها نعمل جدولاً يتضمن ثلاث خانات، الاولى للدرجات، والثانية لإفراغ البيانات، والثالثه للتكرارات.

الجدول (9-2)

التوزيع التكراري لبيانات المثال (2)

الدرجات	إفراغ البيانات	التكرار
0	١	1
1	١	1
2	١	1
3	١١	2
4	١١	2
5	١١	2
6	١١١١	4
7	١١١	3
8	~~١١١١~~	5
9	١١١	3
10	١	1
المجموع الكلي		25

إن المدى وهو الفرق بين اصغر قيمة واكبرها مضافاً له 1، في البيانات السابقة كان 11، وهو يقع ضمن العدد المناسب للفئات في التوزيع التكراري، أما إذا كـان المدى كبيراً 20 أو 50 او 100 فئـة، فإنه يجب أن يتم تقسـيم البيانات الى فئات يتراوح عددها كما اشرنا سابقاً ما بين 5-15 فئة.

مثال (3):

لنفترض أن أحد المعلمين اختبر طلابه في مادة الجغرافيه، وكانت درجاتهم على النحو الآتي:

19	5	1	19	12	4	15	6	12	8
10	16	13	11	8	17	15	20	10	6
16	17	6	14	5	9	17	18	9	11
12	17	9	10	11	14	2	18	13	15
7	12	22	15	21	19	17	8	6	13

المطلوب عرض الدرجات في توزيع تكراري .

الحل: لعرض الدرجات في توزيع تكراري نقوم بما يلي:

1- نحسب المدى= اكبر قيمة – أصغر قيمة + 1

= 22 - 1 + 1 = 22

بما أن المدى أكثر من 15، إذن يجب أن يكون التوزيع ذا فئات.

2- نحدد عدد الفئات Intervals، وعادة ما يتراوح هذا العدد بين 5 - 15 فئة.

ونفترض ان عدد الفئات 11

3- نحدد طول الفئة Interval Width = $\frac{\text{المدى}}{\text{عدد الفئات}} = \frac{22}{11} = 2$

وعندما يكون طول الفئة كسراً، نقرب النتيجة دائماً نحو الاعلى.

4- نعين الحد الادنى للفئة الاولى، ويجب أن يكون هذا الحد مساوياً أو أصغر من اقل قيمة في البيانات وهو 1 .

ونعين الحد الادنى افعلي = الحد الادنى- 0.5

= 1 - 0.5 = 0.5

5- نعين الى الحد الاعلى للفئة الاولى= الحد الادنى الفعلي+ طول الفئة

= 0.5 + 2 = 2.5

ثم نعين الحد الأعلى للفئة الاولى= الحد الاعلى الفعلي- 0.5

= 2.5 - 0.5 = 2

م	حدود الفئة	الحدود الفعلية للفئة
1	2 - 1	2.5 - 0.5

6- نعين الحدود الدنيا والعليا لجميع الفئات الباقية، وذلك بإضافة طول الفئة لكل وحدة، ثم نعين الحدود الدنيا والعليا الفعلية.

7- نعين مراكز الفئات Mid- Point بقسمة حاصل جمع الحد الادنى والاعلى على 2 .

مركز الفئة الاولى= $\frac{2 + 1}{2} = 1.5$

8- نفرغ البيانات المعطاة لدينا على الفئات التي أنشأناها، وذلك بوضع خط عمودي لكل قراءة، وخط مائل للقراءة الخامسة في كل فئة، وذلك لتسهيل جمع التكرارات.

9- نجمع التكرارات المقابلة لكل فئة، ونسجله في عمود التكرارات.

الجدول (9-3) : التوزيع التكراري لبيانات مثال (3)

	حدود الفئة	الحدود الفعلية للفئة	مركز الفئة	افراغ البيانات	التكرار
1	2 - 1	2.5 - 0.05	1.5	ǀǀ	2
2	4 - 3	4.5 - 2.5	3.5	ǀ	1
3	6 - 5	6.5 - 4.5	5.5	~~ǀǀǀǀ~~	6
4	8 - 7	8.5 - 6.5	7.5	ǀǀǀǀ	4
5	10 - 9	10.5 - 8.5	9.5	ǀ ~~ǀǀǀǀ~~	6
6	12 - 11	12.5 - 10.5	11.5	ǀǀ ~~ǀǀǀǀ~~	7
7	14 - 13	14.5 - 12.5	13.5	~~ǀǀǀǀ~~	5
8	16 - 15	16.5 - 14.5	15.5	ǀ ~~ǀǀǀǀ~~	6
9	18 - 17	18.5 - 16.5	17.5	ǀǀ ~~ǀǀǀǀ~~	7
10	20 - 19	20.5 - 18.5	19.5	ǀǀǀǀ	4
11	22 - 21	22.5 - 20.5	21.5	ǀǀ	2
المجموع الكلي					50

طرائق التمثيل البياني للتوزيعات التكرارية:

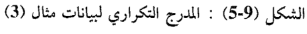
إن الهدف من عرض التوزيعات التكرارية بيانياً هو تسهيل استيعابها والمقارنة بين مفرداتها بطريقة توضح خصائصها بنحو اسهل من التوزيع التكراري.

وهناك طرائق عديدة لعرض البيانات الخاصة بالتوزيعات التكرارية من أهمها وأكثرها استخداماً ثلاث طرائق هي: المدرج التكراري والمضلع التكراري والمنحنى التكراري.

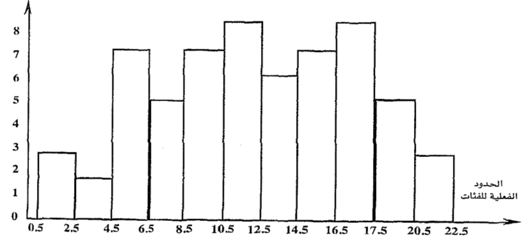
أ - المدرج التكراري Frequency Histogram

يتم في المدرج التكراري تمثيل كل فئة من فئات التوزيع التكراري بعمود (أو مستطيل)، قاعدته (عرض المستطيل) تساوي الحدود الفعلية لتلك الفئة، وارتفاعه يتناسب وتكرارها. وهو قريب الشبه من الأعمدة البيانية، غير أنه في المدرج التكراري تتصل الأعمدة بعضها ببعض، بحيث تكون محصورة بين الحد الأدنى الفعلي والحد الأعلى الفعلي لكل فئة. ولتوضيح كيفية عمل المدرج التكراري نعرض بيانات مثال (3) في الشكل (6).

في البداية نرسم محورين متعامدين، نرصد على المحور السيني (الأفقي) الحدود الفعلية للفئات، وعلى المحور الصادي تكرارات الفئات، وهنا يجب أن يتناسب تقسيم كل محور مع عدد الفئات، أو مع التكرارات كي يظهر الرسم بنحو متناسق، بعدها نقيم على حدود كل فئة عموداً يتناسب ارتفاعه وتكرارات الفئة. والشكل (5-9) يبين ذلك.

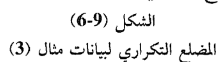


الشكل (5-9) : المدرج التكراري لبيانات مثال (3)



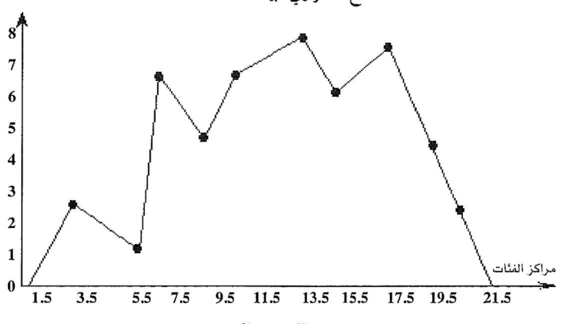
ب - المضلع التكراري Frequency Pollygon

لا تختلف أهداف المضلع التكراري وطريقة عمله عن المدرج التكراري، سوى أننا نتعامل مع مراكز الفئات وليس مع حدودها الفعلية كما حصل في المدرج التكراري، وأما إذا عملنا المدرج ونريد عمل المضلع، فنحدد منتصف الأضلاع العلوية للاعمدة، ونوصل النقاط بعضها ببعض. ويشترط في المضلع التكراري أن نوصل أطرافه (أول نقطة وآخر نقطة) مع مركز فئة افتراضي تتم اضافته قبل التوزيع وبعده حتى نغلق الخط المنكسر. ولتوضيح ذلك نعرض بيانات مثال (3) في مضلع تكراري.



الشكل (6-9)

المضلع التكراري لبيانات مثال (3)

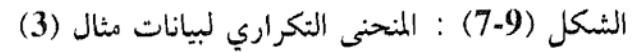


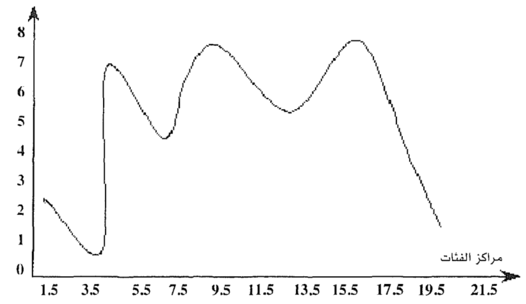
ج- المنحنى التكراري Curve Frequency

يشبه المنحنى التكراري الخط المنحني الذي تم عرضه في طريقة عرض البيانات الأولية.

نحصل على المنحنى التكراري بتمهيد خطوط المضلع التكراري المنكسرة، ولا يشترط أن يمر المنحنى بجميع رؤوس المضلع. أما إذا أردنا عرض بيانات التوزيع في منحنى تكراري فيتم تحديد مراكز الفئات على المحور الأفقي

والتكرارات على المحور العمودي، وبعد تحديد النقاط المقابلة لمركز كل فئة مع التكرار الدال عليها، نوصل بين تلك النقاط باليد مع تمهيد الزوايا، لنحصل على المنحنى التكراري، والشكل (9-7) يبين نموذج للمنحنى التكراري مأخوذ من بيانات مثال (3).

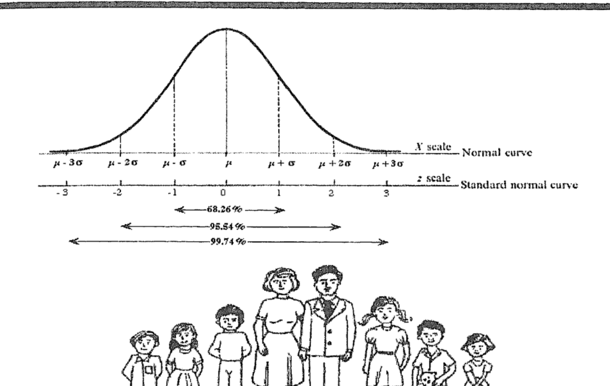

الشكل (9-7) : المنحنى التكراري لبيانات مثال (3)



يستخدم المنحنى التكراري عندما يكون حجم العينة (عدد الطلبة) كبيراً، أكثر من استخدامه في العينات الصغيرة، ومن الأمثلة على العينات الكبيرة، عدد طلبة المدرسة، أو عدد طلبة الصف الأول في مدينة معينة. . . . الخ. وأما العينات الصغيرة فيقصد بها طلبة الصف أو مجموعة محددة من الطلبة.

أشكال المنحنيات Types of Curves

ان لشكل المنحنى التكراري مدلولاً خاصاً ومهماً، إذ من خلاله نستدل على مستوى الطلبة، أو على مدى توزيع ظاهرة معينة أو سمة معينة بين الأفراد، وكذلك نستطيع الحكم على مدى تمثيل العينة للمجتمع. وتميل أغلب الظواهر أو السمات النفسية والتربوية كالتحصيل الدراسي والذكاء وخصائص الشخصية وكذلك الطول والوزن وغيرها الى أن يكون شكل المنحنى فيها قريباً من الجرس المقلوب. والشكل (9-8) يبين نموذج للمنحنى القياسي.



A normal distribution—almost!

الشكل (8-9) المنحنى القياسي وخصائصه الإحصائية

إن النموذج المبين في الشكل (8-9) يسمى منحنى التوزيع الإعتدالي أو الطبيعي Normal Distribution أو المنحنى النموذجي- القياسي، لأن الطبيعة تتوزع فيها الظواهر على أساس الفروق الفردية، أي اختلاف الأفراد في درجة امتلاكهم للسمات، واختلاف الطلبة في مستوى التحصيل، إذ نجد فيهم المتفوق والمتوسط والمتأخر، وللتوزيع الطبيعي خواص عديدة منها:

1 - التوزيع الطبيعي متماثل Symmetrical حول العمود المقام على الوسط الحسابي، وشكله يشبه شكل الجرس المقلوب.

2 - للتوزيع الطبيعي قمة واحدة، أي له منوال واحد فقط (سيتم توضيح مفهوم المنوال في مقاييس النزعة المركزية).

3 - مجموع مساحة التوزيع يساوي 1 .

هذا ومن الناحية العملية، تأخذ المنحنيات أشكالاً متعددة، وذلك بحسب درجة التواء وتفرطح بياناتها.

فالإلتواء SKewness يشير الى مدى بُعد التوزيع عن التماثل وهو يعتمد على درجة تركيز التكرارات عند القيم المختلفة للتوزيع.

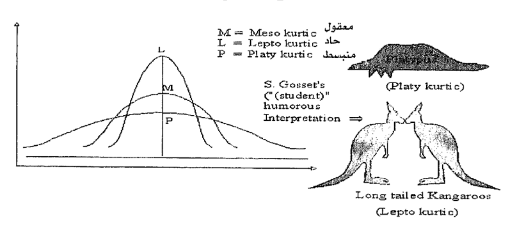
وعندما يكون التوزيع ملتوياً، فله احتمالان، إما أن يكون سالباً Nega-tively SKewed؛ وفيه تكون غالبية الدرجات متجمعة حول النهاية العليا، أي عندما يكون الامتحان سهلاً ويحصل اغلب الطلبة على درجات عالية، وإما أن يكون الإلتواء موجباً Positively Skewed، وفيه تكون غالبية الدرجات متجمعة في النهاية الصغرى، أي عندما يكون الامتحان صعباً وأغلب الطلبة راسبين. والشكلان (9-9) و(9-10) يبينان ذلك.



الشكل (9-9)
منحنى سالب الإلتواء

الشكل (9-10)
منحنى موجب الإلتواء

أما التفرطح (Kurtosis أو التفلطح) فيشير الى مدى تركيز التكرارات في وسط التوزيع، وفي ضوئه تكون التوزيعات مدببة Leptokurtic، أي يتركز عدد كبير من التكرارات قرب الوسط الحسابي، أو تكون التوزيعات مفرطحة Platykurtic، أي عندما يتراكم عدد قليل من التكرارات قرب الوسط. والشكل (9-11) يبين أشكال تفرطح التوزيع.



الشكل (9-11) اشكال تفرطح التوزيع التكراري

مقاييس النزعة المركزية Measures of Central Lendency

يقصد بالنزعة المركزية هو نزوع أو ميل درجات الطلبة الى التركز أو التجمع في وسط التوزيع.

وفي مواقف عديدة نحتاج الى معرفة النزعة المركزية لدرجات مجموعة من الطلبة، أو معرفة اعدادهم في الصفوف، وذلك بعد أن نقوم بعرضها بيانياً سواء أكانت بيانات أولية، أو توزيعات تكرارية. لأن طرائق عرض البيانات غير كافية لإعطاء استنتاج أو وصف دقيق للبيانات، ففي احيان كثيرة تتشابة المنحنيات في توزيعين أو أكثر، أو تكون البيانات المعروضة في جدول ما كثيرة بحيث يصعب ادراكها، في هذه الحالات نلجأ الى معرفة النزعة المركزية للدرجات لنتمكن من اعطاء استنتاجات دقيقة للدرجات.

تقاس النزعة المركزية بمقاييس عديدة من اشهرها ثلاثة هي المتوسط الحسابي، والوسيط، والمنوال، وهناك مقياسان آخران يفيدان في مجال الاحصاء الرياضي هما المتوسط الهندسي، والمتوسط التوافقي. وفيما يلي عرض مفصل للمقاييس الثلاثة الاولى.

1 - المتوسط الحسابي Arithematic Mean

يُعرّف المتوسط الحسابي لمجموعة من الدرجات بأنه: حاصل قسمة مجموع الدرجات على عددها؛ أي أن المتوسط هو القيمة التي لو وزعت على كل فرد من أفراد العينة، لكان مجموع هذه القيم هو المجموع الفعلي للدرجات جميعها. ويرمز للمتوسط الحسابي باللغة العربية الرمز (سَ) ويُقرأ (سين شرطه)، أما في اللغة الانجلزية فله رمزان، الأول $\overline{x}$ (ويقرأ أكس بار) ويستخدم مع العينة، والثاني (M يُقرأ ميو) ويستخدم مع المجتمع.

وتختلف طرائق حساب المتوسط الحسابي تبعاً لطريقة تبويب البيانات العددية ونوعها وكما يلي :

أ- حساب المتوسط من الدرجات الخام

تستخدم هذه الطريقة مع البيانات الأولية التي لا تعرض في توزيع تكراري، وعادة يكون عددها قليلاً (لا يتجاوز 15 درجة).

مثال (4):

لنفترض أن أحد المعلمين اختبر خمسة طلاب في مادة الرياضيات، وحصلوا على الدرجات الاتية:

7 5 6 9 8

فما هو المتوسط الحسابي؟

$$\text{المتوسط الحسابي} = \frac{\text{مجموع الدرجات}}{\text{عدد الطلاب}} \quad \left[\overline{\text{س}} = \frac{\text{مجـ س}}{\text{ن}} \right]$$

$$\overline{\text{س}} = \frac{7+5+6+9+8}{5} = \frac{35}{5} = 7$$

ب- حساب المتوسط من توزيع تكراري بدون فئات

في حالة عرض البيانات في توزيع تكراري طول فئاته 1، ونحتاج أن نحسب متوسط الدرجات لهذا التوزيع، نستخدم المعادلة الآتية:

$$\text{المتوسط} = \frac{\text{مجموع (الدرجة × التكرار)}}{\text{عدد الطلاب}} \quad \left[\overline{\text{س}} = \frac{\text{مجـ (س×ت)}}{\text{ن}} \right]$$

وبالرجوع الى مثال (4)، نلاحظ أن التوزيع التكراري المعروض في جدول (2) يتكون من الدرجات وتكراراتها. وعند تطبيق المعادلة الاحصائية، نضيف عموداً جديداً يتضمن الدرجة مضروبة بالتكرار. والجدول (9-4) يتضمن ذلك.

الجدول (9-4)

الدرجات وتكراراتها لبيانات مثال (4)

الدرجات (س)	التكرار (ت)	الدرجة × التكرار (س×ت)
صفر	1	صفر × 1 = صفر
1	1	1 = 1 × 1
2	1	2 = 1 × 2
3	2	6 = 2 × 3
4	2	8 = 2 × 4
5	2	10 = 2 × 5
6	4	24 = 4 × 6
7	3	21 = 3 × 7
8	4	32 = 4 × 8
9	4	36 = 4 × 9
10	1	10 = 1 × 10
المجموع	25	150

وفي حالة تطبيق المعادلة مباشرة دون اضافة عمود للجدول، تكون كما يأتي:

$$\frac{(\text{صفر} \times 1) + (1 \times 1) + (1 \times 2) \ldots (1 \times 10)}{25} = \frac{150}{25} = 6$$

ج- حساب المتوسط من توزيع تكراري ذي فئات:

عندما تكون البيانات كبيرة نلجأ الى طريقة التوزيع التكراري ذي الفئات (وفيها يكون طول الفئة 2 أو أكثر) لعرض تلك البيانات.

في هذه الحالة، نتعامل مع مركز الفئة والتكرارات في حساب المتوسط الحسابي، باستخدام المعادلة الاتية:

$$\text{المتوسط الحسابي} = \frac{\text{مجموع (مركز الفئة × التكرار)}}{\text{عدد الطلاب}} \quad \left[\bar{\text{س}} = \frac{\text{مجـ (م ف×ت)}}{\text{ن}}\right]$$

مثال:

نطبق هذه المعادلة على بيانات مثال (3) المعروضة في جدول (9-3) فتكون كما يأتي:

$$\text{س} = \frac{(2 \times 1.5) + (1 \times 3.5) \ldots\ldots (21 \times 21.5)}{50}$$

$$= \frac{611}{50} = 12.22$$

د- المتوسط الموزون (متوسط المتوسطات) **Weighted Mean**

عندما تتوافر لدينا متوسطات حسابية لمجاميع عديدة من الطلاب، وأعداد الطلاب في كل مجموعة معلوم، فإن المتوسط العام لتلك المتوسطات يسمى متوسط المتوسطات أو المتوسط الموزون. والمعادلة الاتية تبين طريقة الحساب:

$$\text{متوسط المتوسطات} = \frac{\text{مجموع (متوسط المجموعة × عددها) + (متوسط المجموعة × عددها) + . . .}}{\text{(عدد أفراد مجموعة 1 + عدد أفراد مجموعة 2 + ...)}}$$

$$\frac{\text{مجـ}(\bar{\text{س}}_1 \times \text{ن}_1) + (\bar{\text{س}}_2 \times \text{ن}_2) + \ldots}{\text{ن}_1 + \text{ن}_2 + \ldots\ldots}$$

مثال (5):

اختبر احد المعلمين ثلاث شعب من الطلبة في مادة الكيمياء، وبعد أن قام بحساب متوسط كل شعبة، اراد معرفة المتوسط الحسابي لجميع الشعب.

الشعبة	المتوسط	العدد
أ	85	30
ب	70	33
ج	75	28

$$\text{المتوسط الحسابي} = \frac{(28 \times 75) + (33 \times 70) + (30 \times 85)}{28 + 33 + 30}$$

$$= \frac{6960}{91} = 76.48$$

خواص المتوسط الحسابي

1 - يتأثر المتوسط الحسابي بالقيم المتطرفة، وهذه الخاصية من أهم عيوبه.

2 - يمكن جمع المتوسطات لعدة مجموعات وطرحها في حالة تساوي أعدادها.

3- يتأثر المتوسط بعدد الدرجات، ويميل الى الاستقرار كلما كان العدد كبيراً.

4- يدخل في حساب المتوسط جميع القيم المطلوب حساب المتوسط لها.

5- لا يمكن حساب المتوسط في حالة الجداول المفتوحة Open Tables، سواء من أحد طرفي الجداول، أو من كليهما، بسبب تعذر تحديد مركز الفئة المفتوحة Open Mid Interval.

2 - الوسيط Median

ويسمى بالأوسط، ويُعّرف بأنه الدرجة التي تقع في وسط التوزيع أو منتصفه، أو هو الدرجة التي تقسم توزيع الدرجات على مجموعتين متساويتين من حيث الحجم.

ويرمز له باللغة العربية (و)، وباللغة الانجليزية (Md).

وهناك طرائق عديدة في حساب الوسيط، يتوقف استخدام طريقة معينة دون سواها على نوع البيانات وكمها، وفيما يلي عرض لطرائق حساب الوسيط.

أ- حساب الوسيط للبيانات الاولية

عندما يكون عدد الدرجات قليلاً، لا يستدعي عرضها في جدول أو توزيعها تكرارياً، ونريد حساب وسيطها، نرتبها ترتيباً تصاعدياً أو تنازلياً، ونعدّها، فإذا كان عددها فردياً فإن =

$$\text{رتبة الوسيط} = \frac{\text{عدد الدرجات} + 1}{2} = \left[\text{و} = \frac{\text{ن} + 1}{2} \right]$$

أما إذا كان عدد الدرجات زوجياً، فإن الوسيط هو متوسط القيمتين اللتين رتبهما:

$$\text{رتبة القيمة الأولى} = \frac{\text{عدد الدرجات}}{2} = \left[\frac{\text{ن}}{2} \right]$$

$$\text{رتبة القيمة الثانية} = \frac{\text{عدد الدرجات}}{2} + 1 = \left[\frac{\text{ن}}{2} + 1 \right]$$

مثال (6):

احسب الوسيط لدرجات خمسة طلاب في امتحان الفيزياء.

85 90 60 74 52

الحل:

1 - نرتب الدرجات تصاعدياً :

52 60 74 85 90

2- نحدد رتبة الوسيط: $\frac{\text{ن} + 1}{2} = \frac{5 + 1}{2} = 3$

3- الدرجة 74 هي الوسيط؛ لأن رتبتها الثالثة.

مثال (7): احسب الوسيط لدرجات ستة طلاب في امتحان الكيمياء:

88 68 54 70 95 67

الحل:

1- نرتب الدرجات تصاعدياً: 54 67 68 70 88 95

2- نحدد رتبة القيمة الوسيطية الأولى: $= \frac{ن}{2} = \frac{6}{2} = 3$

3- نحدد رتبة القيمة الوسيطية الثانية: $= \frac{ن}{2} + 1 = \frac{6}{2} + 1 = 4$

4- القيمة الوسيطية الأولى 68، والقيمة الوسيطية الثانية 70

5- متوسط القيمتين الوسيطيتين: $\frac{70 + 68}{2} = 69$

ب - حساب الوسيط للتوزيعات التكرارية

عندما يتم عرض الدرجات في توزيع تكراري سواء أكانت بفئات أم بدون فئات، فإن الوسيط يحسب على وفق المعادلة الآتية:

$$\text{الوسيط} = \text{الحد الفعلي الأول لفئة الوسيط} + \frac{\frac{\text{عدد الدرجات}}{2} - \text{التكرار المتجمع السابق لفئة الوسيط}}{\text{تكرار فئة الوسيط}} \times \text{طول الفئة}$$

$$و = ل + \frac{\frac{ن}{2} - ت\ ق}{ت} \times ض$$

مثال: وبالعودة الى بيانات مثال (3) الوارد ضمن موضوع التوزيع التكراري ذي الفئات نستخرج وسيط درجاته.

الحل:

1- نضيف عمود جديد بعد عمود التكرار، يسمى التكرار المتجمع الصامد، والتكرار المتجمع Cumulative Frequency لفئة ما هو

مجموع تكرارات جميع القيم التي تساوي أو تقل عن الحد الأعلى الفعلي لتلك الفئة.

ولحساب هذا التكرار، نبدأ بالفئة الاولى، فيكون تكرارها المتجمع هو تكرارها الفعلي، أما الفئة الثانية، فإن تكرارها المتجمع هو تكرارها مضافاً إليه تكرار الفئة الاولى، والفئة الثالثة تكرارها المتجمع هو تكرارها الفعلي مضافاً إليه التكرار المتجمع الصاعد السابق، وهكذا. والجدول (9-5) يبين ذلك.

الجدول (9-5) : فئات الدرجات والتكرار المتجمع لبيانات مثال (3)

م	حدود الفئة	الحدود الفعلية للفئة	التكرار	المتجمع الصاعد
1	2-1	2.5-0.5	2	2
2	4-3	4.5-2.5	1	3
3	6-5	6.5-4.5	6	9
4	8-7	8.5-6.5	4	13
5	10-9	10.5-8.5	6	19
6	12-11	12.5-10.5	7	26
7	14-13	14.5-12.5	5	31
8	16-15	16.5-14.5	6	37
9	18-17	18.5-16.5	7	44
10	20-19	20.5-18.5	4	48
11	22-21	22.5-20.5	2	50

2- نحدد رتبة الفئة الوسيطية = $\frac{ن}{2} = \frac{50}{2} = 25$

3- نبحث في التكرار المتجمع الصاعد عن التكرار 25، وفي حالة عدم وجود نفس التكرار، نأخذ التكرار المتجمع الأعلى منه مباشرة، ونعده تكرار الفئة الوسيطية. وفي مثالنا هذا تكون الفئة السادسة هي الفئة الوسيطية.

4- نطبق معادلة حساب الوسيط:

$$و = 10.5 + \frac{19 - \frac{50}{2}}{7} \times 2$$

$$= 10.5 + \frac{6}{7} \times 2 = 12.21 \text{ الوسيط}$$

خواص الوسيط

1- يتأثر الوسيط بالدرجات الوسطى أكثر من تأثره بالدرجات المتطرفة في التوزيع التكراري.

2- مجموع الإنحرافات المطلقة عن الوسيط، أصغر من مجموع الانحرافات المطلقة عن المتوسط الحسابي.

مثال :

2 2 3 5 8

المتوسط الحسابي= 4 مجموع الانحرافات المطلقة = 10

الوسيط = 3 مجموع الانحرافات المطلقة = 9

3- الوسيط احد المقاييس الرتبية، وعلى هذا فهو مقياس مناسب للظواهر الكمية الرتبية.

4- يعاب عليه عدم استخدامه لكل البيانات المتاحة عند حسابه.

5- يمكن ايجاده في حالة الجداول التكرارية المفتوحة.

3 - المنوال Mode

يشير المنوال الى ما هو شائع أو دارج بين الافراد، او هو الدرجة الاكثر شيوعاً او تكراراً من غيرها من الدرجات.

ويرمز له باللغة العربية (م)، وباللغة الانجليزية (Mo).

مثال (8):

اذا كانت درجات مجموعة من الطلاب في امتحان معين هي:

8 9 6 7 8

فما هو المنوال؟

الحل:

المنوال هو 8، لأنه تكرر مرتين، وبذلك هو الدرجة الاكثر تكراراً في هذه البيانات. أما إذا لم يكن هناك درجة يقابلها تكرار اكبر مما يقابل غيرها؟ فلا يكون هناك منوال.

وفي بعض الاحيان يكون اكثر من منوال واحد للتوزيع Multi- Modal Distribution، فلو أضفنا الدرجة (9) لبيانات المثال السابق، لأصبح له منوالان Bi- Modal Distridution هما: 8 و9؛ لانهما تمتلكان التكرار نفسه وهو 2.

وفي حالة التوزيعات التكرارية البسيطة بدون فئات، فإن المنوال هو الدرجة المقابلة لأكبر تكرار. أما اذا كانت للتوزيعات التكرارية فئات، فإن المنوال يحسب على وفق المعادلة الاتية:

$$\text{المنوال} = \text{الحد الأدنى للفئة المنوالية} + \frac{ف_1}{ف_1 + ف_2} \times \text{طول الفئة المنوالية}$$

حيث أن:

$ف_1$ = الفرق بين تكرار الفئة المنوالية والتكرار السابق.

$ف_2$ = الفرق بين تكرار الفئة المنوالية والتكرار اللاحق.

مثال (9): احسب المنوال للتوزيع التكراري الآتي:

حدود الفئة	التكرار	الفروق
59 - 55	7	$ف_1 =$
64 - 60	15	$ف_2 =$
69 - 65	8	

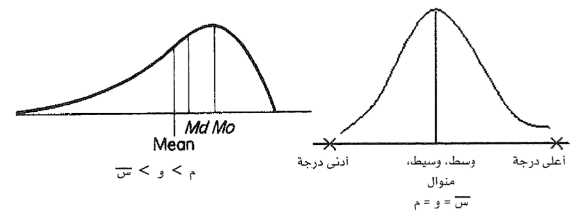
خواص المنوال

1- يتأثر المنوال بعدد فئات التوزيع وبطول الفئة.
2- عندما تتعدد قمم التوزيع تتعدد أيضاً قيم المنوال.
3- يمكن استخراج المنوال في الجداول المفتوحة، بشرط ألا يكون اكبر تكرار أمام هذه الفئة.

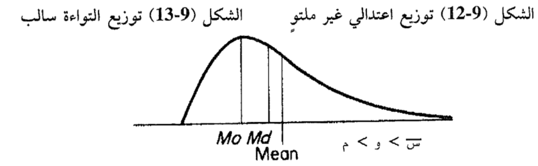
العلاقة بين مقاييس النزعة المركزية

ثمة علاقة قوية بين مقاييس النزعة المركزية؛ وتتوقف قوة تلك العلاقة على طبيعة توزيع البيانات.

ففي حالة التوزيعات المتماثلة (التوزيع الاعتدالي)، يكون للوسط الحسابي والوسيط والمنوال نفس القيمة؛ أما إذا كان التوزيع ملتوياً سواء أكان التواءً سالباً أم موجباً، فإن قيم النزعة المركزية تختلف. والاشكال الآتية (9-12، 9-13، 9-14) تبين ذلك.



الشكل (9-12) توزيع اعتدالي غير ملتوٍ   الشكل (9-13) توزيع التواءة سالب



الشكل (9-14) توزيع التواءه موجب

وفي حالة المنحنيات التكرارية بسيطة الالتواء، وجد عالم الاحصاء بيرسون C. Pearson بعد أن أجرى الكثير من المحاولات، ان مقاييس النزعة المركزية الثلاثة تربط بينها علاقة تبينها المعادلة الآتية:

س̄ - م = 3 (س̄ - و) أي أن: س̄ = 0.5 (3و - م)

العمليات الأربع ومقاييس النزعة المركزية

تتأثر مقاييس النزعة المركزية الثلاثة بمقدار القيمة وبنوع العملية التي تجرى على القيم الأصلية. فلو اضفنا قيمة ثابتة الى جميع القيم فإن الوسط والوسيط والمنوال سيتأثران بنفس القيمة المضافة وهكذا عندما نقوم باستخدام عملية الطرح والضرب والقسمة.

فلو كانت لدينا القيم (8، 10، 3، 2، 2) فإن الوسط هو 5، والوسيط 3، والمنوال 2. عند اضافة 3 لجميع القيم سيكون الوسط 8، والوسيط 6، والمنوال 5. أما اذا قسمنا القيم الأصلية على 2، فإن المتوسط سيكون 2.5، والوسيط 1.5، والمنوال 1. وهكذا مع عملية الطرح والضرب.

هذه الخاصة المهمة لمقاييس النزعة المركزية يمكن الافادة منها عند تحويل درجات الطلبة في امتحان ما من مقياس مئوي الى مقياس عُشري؛ فإذا كان المتوسط في المقياس المئوي 80% فإن المتوسط في المقياس العشري هو 8 من 10.

علاقة مقاييس النزعة المركزية بمستويات القياس

لكل مستوى من مستويات القياس الأربعة خاصية تميزه عن غيره من المستويات، وطبقاً لهرمية مستويات القياس فإن القياس النسبي تتوافر فيه أربع خصائص والقياس الفئوي ثلاث خصائص وهكذا وصولاً للقياس الإسمي فيه خاصية واحدة فقط وهي التصنيف.

وعلى أساس مستوى القياس يمكن استخدام مقياس النزعة المركزية المناسب، فالمنوال ينفع مع جميع مستويات القياس، والوسيط ينفع مع القياس الرتبي والفئوي والنسبي، أما المتوسط فينفع مع القياس الفئوي والنسبي.

وفيما يلي توضيح لذلك:

مستوى القياس	مقياس النزعة المركزية المناسب
قياس اسمي	منوال
قياس رتبي	منوال+ وسيط
قياس فئوي	منوال+ وسيط+ متوسط
قياس نسبي	منوال+ وسيط+ متوسط

وطبقاً لجلاس وهوبكنز (Glass & Hopkins, 1996) فإن المتوسط الحسابي هو أكثر مقاييس النزعة المركزية استقراراً وثباتاً reliable (لأن مجموع مربعات انحرافات القيم عن المتوسط الحسابي أقل من الوسيط ومن المنوال)، لذلك نجد المتوسط الحسابي هو الاكثر استخداماً في البحوث والتقارير العلمية.

مقاييس التشتت Variability Measures

تفيد مقاييس النزعة المركزية (الوسط، الوسيط، المنوال) في وصف التوزيعات التكرارية دون تحديد لصفاتها؛ فربما يكون لدينا ظاهرتان متساويتان في أحد مقاييس النزعة المركزية كالمتوسط مثلاً، إلا أنهما مختلفتان في الدرجات.

مثال (10)

لو كانت درجات مجموعة من الطلاب في امتحان اللغة العربية هي:

26 30 23 35 21 28 19 فالمتوسط هو= 26

واذا كانت درجات مجموعة أخرى من الطلاب في الامتحان نفسه هي:

17 33 14 42 20 30 26 فالمتوسط هو: 26

هذا يعني أن متوسط الدرجات في المجموعتين متساو، إلا أنه بالنظر الى درجات كل مجموعة نجد أن هناك اختلافاً بينهما، وهذا يعني أن الوسط الحسابي لا يكفي لوصف البيانات، أو للحكم على تشابهها.

إذن لا بد من استعمال مقاييس أخرى تبين لنا مدى اختلاف البيانات فيما بينها، ومن هذه المقاييس مقاييس التشتت.

إن الشيء المهم في قياس تشتت الدرجات هو معرفة مدى انحرافها بالنسبة إلى بعضها البعض، وبالنسبة الى القيمة المتوسطة.

ولقياس التشتت وسائل عديدة من اكثرها استخداماً في المجالات التربوية والنفسية المدى، والتباين، والانحراف المعياري، فضلاً عن الانحراف الربيعي، والانحراف المتوسط، والتشتت النسبي. سيتم التعرض إلى الوسائل الثلاث الأولى.

(1) المدى Range

يُعرّف المدى (في حالة البيانات الأولية) بأنه الفرق بين أكبر قيمة وأصغر قيمة مضافاً اليه وحدة دقة واحدة، ويُعرّف (في حالة التوزيع التكراري) بأنه الفرق بين الحد الأعلى الفعلي للفئة العليا والحد الادنى الفعلي للفئة الدنيا. وبشكل عام، يُعرّف المدى بأنه الفرق بين اكبر قيمة فعلية واقل قيمة فعلية في التوزيع.

يفيد المدى إذا تم استخدامه جنباً الى جنب مع أحد مقاييس النزعة المركزية للمقارنة بين توزيعين أو اكثر من البيانات، أو عند الحكم على كفاءة شخص لشغل وظيفة أو مهنة معينة، فعندما يتقدم شخصان لشغل وظيفة مدرب، ومتوسط درجاتهما متساو فإننا نلجأ إلى المدى في عملية الاختيار، فالذي يكون مداه صغيراً هو الذي سيقع عليه الاختيار.

مثال (11)

إذا كانت درجات مجموعة من الطلبة هي:

69 70 98 77 82 56 47 78 20

فإن المدى= اكبر درجة- اصغر درجة + 1

$= 98 - 20 + 1 = 79$

ومن خلال المثال السابق، نستنتج بأن المدى لا يعتمد على جميع الدرجات، بل على درجتين فقط (أكبر درجة وأصغرها)، وهذا يقلل من اهميته، حيث أنه ربما تكون هاتان الدرجتان شاذتين (متطرفتين).

وفي حالة المثال السابق، نجد أن معظم الدرجات تقع بين 47- 82، أي انها متقاربة من بعضها اكثر من المدى، في هذه الحالة يمكن تحسين المدى، وحذف الدرجات الشاذة، فمدى الدرجات السابقة بعد تحسينه من خلال حذف الدرجات المتطرفة (20، و98) يكون:

المدى= 82- 47+ 1= 36 (وهو أقل بكثير من المدى السابق).

وأما المدى في التوزيعات التكرارية، فمن خلال الرجوع الى البيانات في مثال (3) نجد أن الحد الأدنى الفعلي للفئة الدنيا هو 0.5، والحد الاعلى الفعلي للفئة العليا هو 22.5، وبذلك يكون المدى:

المدى = الحد الاعلى الفعلي للفئة العليا- الحد الادنى الفعلي للفئة الدنيا
= 22.5- 0.5= 22

خصائص المدى

إن المدى سهل الحساب لاعتماده على قيمتين فقط، ولكنه غير دقيق، وعليه فالمدى مفيد في اعطاء فكرة سريعة عن طبيعة توزيع البيانات، وهو مضلل وغير مفيد عندما يستخدم لمقارنة مجموعتين تختلفان في الحجم، فزيادة حجم المجموعة ربما يزيد من احتمالية وجود قيم متطرفه، كما أن المدى غير مفيد عندما يتساوى مدى مجموعتين من البيانات ويختلف متوسطهما الحسابي. ورغم كل ذلك يمكن ان يُعد المدى خطوة أولى في معرفة طبيعة تشتت التوزيع، على أن يتبعه حساب التباين او الانحراف المعياري.

(2) التباين Variance

يُعرّف التباين بأنه متوسط مجموع مربعات انحرافات الدرجات عن متوسطها الحسابي.

ولحساب التباين طرائق عديدة، تختلف تبعاً لنوع البيانات وعددها، ففي حالة البيانات الاولية (غير الموزعة تكرارياً) نستخدم المعادلة الآتية:

$$\text{التباين} = \frac{\text{مجموع (الدرجة - المتوسط الحسابي)}^2}{\text{عدد الدرجات} - 1}$$

$$ع^2 = \frac{\text{مجـ (س} - \bar{\text{س}})^2}{\text{ن} - 1}$$

مثال (12)

احسب التباين للدرجات الآتية:

14 6 11 9 13 15 12 10 7 3

الحل:

1- نحسب المتوسط الحسابي $= \frac{3 + 7 + 10 + \ldots 14}{10 - 1} = \frac{100}{10} = 10$

2- نطبق معادلة حساب التباين:

$$ع^2 = \frac{(3-10)^2 + (7-10)^2 + (10-10)^2 + \ldots (14-10)^2}{9}$$

$$= \frac{130}{9} = 14.44$$

أما في حالة توزيع البيانات توزيعاً تكرارياً بفئات أو بدون فئات، فإننا نحسب التباين على وفق المعادلة الآتية:

$$\text{التباين} = \frac{\text{مجموع (الدرجة - المتوسط الحسابي)}^2 \times \text{التكرار}}{\text{عدد الدرجات} - 1}$$

$$ع^2 = \frac{\text{مجـ (س} - \bar{\text{س}})^2 \times \text{ت}}{\text{ن} - 1}$$

هذه المعادلة تناسب التوزيعات بدون فئات، وفي حالة الفئات نعوض الدرجة (س) بمركز الفئة.

مثال (13)

احسب التباين للتوزيع التكراري الآتي:

الجدول (9-6) الفئات وتكراراتها لمثال افتراضي

م	حدود الفئة	التكرار	مركز الفئة م ف	(م ف - سَ)^2	(م ف - سَ)^2× ت
1	28 - 24	10	26	169	1690
2	33 - 29	3	31	64	192
3	38 - 34	10	36	9	90
4	43 - 39	8	41	4	32
5	48 - 44	12	46	49	588
6	53 - 49	7	51	144	1008
المجموع		80	المجموع		3600

الحل :

1- نحسب المتوسط الحسابي للبيانات وهو:

$$\bar{\text{س}} = \frac{\text{مجـ(م ف × ت)}}{\text{ن}} = \frac{(10 \times 26) + (3 \times 31) + \ldots (7 \times 51)}{50} = 39$$

2- نضيف عمود ثاني ونطرح فيه مركز الفئة من المتوسط الحسابي، ونربع النتيجة (م ف × ت)^2 .

3- نضيف عمود ثالث، نضرب فيه الناتج السابق في التكرار المقابل لكل فئة (م ف- $\bar{\text{س}}$)^2 × ت.

4- نطبق معادلة حساب التباين وهي:

$$\text{ع}^2 = \frac{\text{مجـ(م ف} - \bar{\text{س}})^2 \times \text{ت}}{\text{ن} - 1} = \frac{3600}{49} = 73.47$$

وبما أن هذه الطريقة تحتاج وقتاً طويلاً، هناك طريقة مختصرة لحساب

التباين تتم باستخدام المعادلة الآتية:

$$ع^2 = \frac{مجـ س^2 - ن \times (\bar{س})^2}{ن - 1}$$

وعند تطبيق هذه المعادلة على البيانات الواردة في مثال (12) نكتب الدرجات في عمود، ومربعاتها في عمود آخر:

س	$س^2$
3	9
7	49
10	100
12	144
15	225
13	169
9	81
11	121
6	36
14	196
مجـ 100	1130

$$ع^2 = \frac{1130 - 10 \times (10)^2}{9} = 14.44$$

وهي النتيجة نفسها التي توصلنا اليها بالطريقة التقليدية.

وفي حالة البيانات الموزعة تكرارياً بفئات او بدون فئات نستخدم المعادلة الآتية:

$$\text{التباين} = \frac{1}{\text{عدد الدرجات} - 1}\left[(\text{مجموع الدرجات مربعة} \times \text{التكرار}) - \text{عدد الدرجات} \times \text{المتوسط}^2\right]$$

$$ع^2 = \frac{1}{ن - 1} [(مجـ س^2 \times ت) - ن \times (\overline{س})^2]$$

وعند تطبيق هذه المعادلة على البيانات الواردة في المثال (13)، تكون خطوات الحل هي:

1- نحسب المتوسط الحسابي.

2- نربع مركز الفئة (م ف)^2

3- نضرب مربع مركز الفئة بالتكرار المقابل لها (م ف)^2 × ت

4- نستخرج حاصل جمع الخطوات السابقة في (م ف)^2 × ت والجدول (9-7) يبين هذه الخطوات.

الجدول (9 - 7)

الفئات وتكراراتها لبيانات المثال (13)

الفئات	التكرار	مركز الفئة	م ف × ت	م ف^2	(م ف)^2 × ت
28 - 24	10	26	260	676	2760
33 - 29	3	31	93	961	2883
38 - 34	10	36	360	1296	12960
43 - 39	8	41	328	1681	13448
48 - 44	12	46	55	2116	25392
53 - 49	7	51	357	2601	18207
المجموع	50		1590		79650

وعند تطبيق المعادلة يكون الناتج :

$$ع^2 = \frac{1}{49} [79650 - 50 \times (39)^2] = 73.47$$

خصائص التباين:

1- يتأثر حساب التباين بكافة القيم في التوزيع.

2- لا يصلح في حالة وجود قيم متطرفة في التوزيع.

3- اكثر مقاييس التشتت استخداماً في عمليات الاحصاء التحليلي.

4- يتأثر التباين بعمليتي الضرب والقسمة للقيم الاصلية، ولا يتأثر بعمليتي الجمع والطرح. فعند اضافة قيمة ثابتة او طرح قيمة ثابتة من جميع القيم الاصلية نجد ان التباين لا يتغير.

5 - يقاس التباين بوحدات مربعة مقارنة بقيم البيانات الأصلية.

(3) الانحراف المعياري Standard Deviation

يُعرّف الانحراف المعياري بأنه الجذر التربيعي الموجب للتباين، والسبب وراء اخذ الجذر التربيعي، لأن التباين هو مربع وحدة البيانات، والانحراف المعياري هو مقياس طبيعي لتشتت الدرجات، وعلى هذا فالإنحراف المعياري هو مقياس تشتت طبيعي للدرجات، بينما التباين هو مقياس تشتت مبالغ فيه، أو غير طبيعي للدرجات. وبذلك يمكن تعريف التباين بأنه مربع الانحراف المعياري.

وعلى اختلاف انواع البيانات (أولية، أو توزيعات تكرارية) يتم حسابها على وفق المعادلة الاتية:

$$\text{الانحراف المعياري} = \sqrt{\text{التباين}}$$

$$\text{ع} = \sqrt{\text{ع}^2}$$

فالدرجات التي تباينها 73.47، انحرافها المعياري 8.57، والتي تباينها 14.44 انحرافها المعياري 3.8 .

خصائص الانحراف المعياري

1- يُعد الانحراف المعياري من اكثر مقاييس التشتت استخداماً وشيوعاً لاعتماده في العديد من المعادلات الإحصائية المتعلقة باختبار الفرضيات وإجراء المقارنات.

2- للإنحراف المعياري قيمة موجبة لا يمكن ان يكون سالباً لانه مقياس مسافة أو بعد.

3- هناك انحراف معياري للمجتمع وآخر للعينة، الفرق بينهما يتمثل في قسمة بيانات المجتمع على(ن)، وقسمة بيانات العينة على (ن-1).

4- يعبر الانحراف المعياري عن مقدار تشتت القيم عن متوسطها الحسابي، فإذا كان الانحراف المعياري 1.5 فهذا يعني ان البيانات تنحرف او تبتعد عن متوسطها بمقدار 1.5 .

5- كما هو حال التباين، يتأثر الانحراف المعياري بكافة قيم التوزيع، ولذلك فهو غير مفيد في حالة وجود قيم متطرفة.

6- يتأثر بعمليتي الضرب والقسمة ولا يتأثر بالجمع والطرح للقيم الاصلية للتوزيع.

العلاقة بين المدى والانحراف المعياري:

ثمة علاقة بين المدى والانحراف المعياري، فقد دلت التجارب على ان نسبة المدى الى الانحراف المعياري تقع بين 2-8 أضعاف للعينات التي يقارب حجمها المائة.

فعندما يكون الإنحراف المعياري 2.61 والمدى 12، فإن ضعف الإنحراف المعياري هو (5.21)، وثمانية أمثاله هي (20.88)، وبما أن المدى 12 فإنه يقع بين 5.21- 20.88. وفي حال وقوع المدى خارج هذه الارقام، يتعين مراجعة طريقة حساب الانحراف المعياري فربما هناك خطأ ما.

مقاييس العلاقة Measures of Relationship

تقدم مقاييس النزعة المركزية ومقاييس التشتت وصفاً إحصائياً لتوزيع درجات الطلبة في مادة دراسية معينة، وفي احيان كثيرة يشد أو يجذب انتباه المعلم وجود علاقة بين درجات الطلبة في مادة الرياضيات مثلاً ودرجاتهم في

الفيـزياء، أو درجاتهم في اللغة العربية واللغة الانجليزية. ولكي يتعرف المعلم طبيعة تلك العلاقة، يحتاج الى استخدام مقاييس احصائية أخرى غير النزعة المركزية والتشتت، تسمى بمقاييس العلاقة.

ومقاييس العلاقة تقدم وصفاً كمياً (أو رقمياً) بين درجات الطلبة في مادة معينة ودرجاتهم في مادة أخرى، أو أي متغيرين توجد بينهما علاقة منطقية. وفي هذه الحالة نحتاج ان نرصد درجتين عن كل طالب، الأولى للمتغير الاول (س)، والثانية للمتغير الثاني (ص). ولتحديد اتجاه الاقتران بين درجات المتغيرين لا بد من حساب معامل الارتباط.

ومعامل الارتباط هو رقم يبين مدى الاقتران أو الترابط او العلاقة بين المتغيرين، ويتراوح عادة معامل الارتباط بين (+1) و (-1) مروراً بالصفر. ويشير الرقم الى قوة العلاقة بين المتغيرين، بينما تشير الاشارة (-، +) الى اتجاهها.

وهناك طرائق عديدة في حساب معامل الارتباط، منها بيرسون وسبيرمان، وبايسيريال و. . . .الخ. وتختلف استخدامات كل طريقة تبعاً لمستوى قياس المتغير (اسمي، رتبي، فئوي، نسبي).

1 - معامل ارتباط بيرسون **Pearson Correlation Coefficient**

وهو من اكثر معاملات الارتباط شيوعاً واستخداماً في مجال القياس والتقويم وخاصة عند حساب ثبات اختبارات التحصيل وصدقها:

يحسب معامل ارتباط بيرسون على وفق المعادلة الآتية:

$$ر = \frac{\text{مجـ س ص} - (\text{مجـ س}) \times (\text{مجـ س})}{\sqrt{[\text{ن مجـ س}^2 - (\text{مجـ س})^2]\ [\text{ن مجـ ص} - (\text{مجـ ص})^2]}}$$

مثال (14)

احسب معامل الارتباط بين درجات خمسة طلاب في مادتي الرياضيات والفيزياء.

الطالب	الرياضيات	الفيزياء
1	12	10
2	14	14
3	8	10
4	14	12
5	12	4

الحل:

1 - نحسب مجموع الدرجات في الرياضيات والفيزياء، ومجموع ضرب درجات الرياضيات والفيزياء، ومجموع الدرجات مربعة في الرياضيات والفيزياء. وعليه نحتاج الى استحداث ثلاثة اعمدة جديدة وبذلك يكون الجدول الاتي:

الجدول (9 - 8) درجات الطلاب في الفيزياء والرياضيات

الطالب	الرياضيات (س)	الفيزياء (ص)	س × ص	$س^2$	$ص^2$
1	12	10	120	144	100
2	14	14	196	196	196
3	8	10	80	64	100
4	14	12	168	196	144
5	12	4	48	144	16
المجموع	60	50	612	744	556

2- نطبق معادلة بيرسون على البيانات السابقة:

$$ر_{2.1} = \frac{(50 \times 60) - 612 \times 5}{\sqrt{\left[(50)^2 - 556 \times 5\right]\left[(60)^2 - 744 \times 5\right]}} = 0.327$$

2 - معامل ارتباط سبيرمان للرتب

Sperman's Rank Correlation Coefficient

ويستخدم لايجاد العلاقة بين متغيرين رتبيين مثال ذلك العلاقة بين رتب مجموعة من الطلبة في التكيف الاجتماعي ورتبهم وفقاً لمستوياتهم العلمية، فهذا المعامل يتعامل مع رتب هذه المجموعة وليس مع درجاتهم الاصلية.

مثال (15)

سألَ معلم أن يرتب مجموعة مكونة من ستة طلاب من صفه على متغيرين الاول وفقاً لتكيفهم الاجتماعي، والثاني وفقاً لمستواهم العلمي، فكان ترتيبهم على المتغيرين كما يأتي:

الطالب	الترتيب في التكيف الإجتماعي
أ	1
ب	2
ج	3
د	4
هـ	5
و	6

الترتيب في التكيف العلمي	ف	$ف^2$
2	1-	1
3	1-	1
5	2-	4
4	0	0
1	4	16
6	0	0
		22

الحل:

$$ر = 1 - \frac{6 \text{ مجـ } ف^2}{ن(ن^2 - 1)}$$

حيث أن ف: الفرق بين ترتيب الطالب على المتغير الاول والمتغير الثاني.

ن: عدد أفراد المجموعة.

1 - نحسب الفرق بين كل رتبة من س والرتبة المقابلة لها في ص. نحصل على (ف).

2 - نربع الفروق بين كل رتبتين (س)و (ص)، ونجمع النتائج لنحصل على ($ف^2$) ومقداره في المثال اعلاه = 22

3 - نطبق معادلة سبيرمان فنحصل على معامل الارتباط.

$$ر= 1 - \frac{22 \times 6}{(1 - 36)\,6}$$

$$= 1 - \frac{22 \times 6}{(1 - 36)\,6} = 1 - 0.628$$

ر= 0.372 معامل الارتباط بين رتب المجموعة على المتغيرين.

3 - معامل ارتباط بايسيريال **Biserial Correlation Coefficient**

ويستخدم هذا المعامل لايجاد العلاقة بين متغيرين، الأول ثنائي متقطع مثل العلامة التي تخصص لكل فقرة في اختبار تحصيلي (صفر، 1)، والمتغير الثاني يكون متصلاً مثل العلامة الكلية التي يحصل عليها الطالب في اختبار معين.

فلو أردنا على سبيل المثال التأكد من نوع وقوة العلاقة بين علامات مجموعة من الطلاب على فقرة معينة وعلاماتهم على الاختبار ككل أي العلاقة بين الفقرة والاختبار من أجل تحديد بقاء تلك الفقرة أو حذفها من الاختبار، فان ذلك يتقرر من خلال معرفة قيمة معامل ارتباط بايسيريال .

مثال (16)

لنفرض أننا نريد إيجاد العلاقة بين الدرجة على الفقرة (1) مع عموم اختبار يتكون من (50) فقرة طبق على مجموعة من الطلاب كما مدرجة في أدناه:

الطلاب	علامة الطالب على الفقرة (1)	علامة الطالب على الاختبار
أ	1	45
ب	0	38
ج	1	49
د	1	40
هـ	0	35
و	1	45
ل	0	42
م	0	44
ن	1	48
ي	1	47

الحل:

$$ر_{بب} = \frac{\overline{ص}_1 - \overline{ص}_0}{ع} \sqrt{ل\ ك}$$

حيث أن: $ر_{بب}$ = قيمة معامل ارتباط بونيت بايسيريال

$\overline{ص}_1$، $\overline{ص}_0$ = متوسط قيم العلامات الكلية للطلبة الحاصلين على (1) و (صفر) على الفقرة رقم (1) على التوالي.

ل، ك = نسبة الطلبة الذين حصلوا على علامة(صفر) ونسبة الطلاب الذين حصلوا على علامة (1) على الفقرة رقم (1) على التوالي.

ع = الانحراف المعياري للدرجات الكلية.

الخطوات :

1 - متوسط درجات الإجابة الصحيحة = 45.67

2 - متوسط درجات الإجابة الخاطئة = 39.75

3 - الإنحراف المعياري للدرجات الكلية = 4.29

4 - نسبة الطلبة الحاصلين على 1 = 0.6

5 - نسبة الطلبة الحاصلية على صفر = 0.4

6 - نطبق المعادلة السابقة الذكر لايجاد معامل ارتباط بونيت بايسيريال

$$= \frac{39.75 - 45.67}{4.29} \sqrt{0.4 \times 0.6}$$

$$= 0.49 \times 1.38$$

$$= 0.676$$

قيمة معامل الارتباط بين علامات المجموعة على الفقرة (1) وعلاماتهم الكلية على الاختبار ويمثل علاقة قوية وموجبة بين المتغيرين.

تفسير معامل الارتباط

هناك مؤشرين للحكم على معامل الارتباط وهما:

1- القوة، وهي درجة معامل الارتباط، وتتراوح بين صفر إلى 1.00، وتشير الى قوة العلاقة بين المتغيرين، فإذا كانت صفراً دلت على انعدام العلاقة.

2- الاتجاه، وهو الإشارة الموجودة الى جنب الرقم، إما ان تكون موجبة أو سالبة، فإذا كانت موجبة دلت على وجود علاقة طردية، بمعنى أن ارتفاع درجات الرياضيات يقابله ارتفاع في درجات الفيزياء، وكلما انخفضت درجات الرياضيات يقابله انخفاض في درجات الفيزياء، أما إذا كانت سالبة فتدل على وجود علاقة عكسية، بمعنى ان ارتفاع درجات الرياضيات يقابله انخفاض درجات الفيزياء أو العكس.

دلالة معامل الارتباط

إن الحكم على قوة الارتباط يعتمد على حجم العينة، وهناك جداول احصائية خاصة بمعامل الارتباط يمكن استخدامها للحكم على مستوى الدلالة.

تتلخص طريقة استخدام الجدول الاحصائية (الجدول،9)، بمقارنة معامل الإرتباط بالقيمة الجدولية اعتماداً على حجم العينة (ن- 2)، ومستوى الدلالة (0.05 أو 0.01)، فإذا كانت القيمة المحسوبة اعلى من الجدولية، يكون الارتباط دالاً إحصائياً، أي توجد علاقة بين المتغيرين، أما اذا كانت القيمة المحسوبة اقل من الجدولية، فهذا يعني أن الارتباط غير دال احصائياً.

وفي المثال السابق، فإن الحكم على الارتباط 0.33، يتطلب تحديد درجات الحرية وهي (5-2= 3)، وتحديد مستوى الدلالة (0.05) أو (0.01)، وفي الجدول نجد أن درجة الحرية (3) تقابلها 0.878 عند مستوى (0.05)، و 0.959 عند مستوى (0.01) وبما ان القيمة المحسوبة أقل من القيمتين الجدوليتين، فإنه غير دال احصائياً. وهذا يعني لا توجد علاقة بين درجات الطلبة في الرياضيات والفيزياء.

أما اذا كانت العلاقة دالة إحصائياً، فيتعين ان ننظر الى الاشارة سالبة أو موجبة، فإذا كانت سالبة دلت على العلاقة العكسية بين المتغيرين، وإذا كانت موجبة دلت على العلاقة الطردية بين المتغيرين.

وهناك طريقة أخرى للحكم على العلاقة بين المتغيرات، تعتمد على القيمة المحسوبة بغض النظر عن حجم العينة، ويشير النبهان (2005) الى المستويات الآتية:

الجدول (9-9) القيم الجدولية لدلالة معامل الارتباط

قيمة الارتباط	الحكم على الارتباط
1.00 (+ أو -)	تام
0.99 - 0.80	عالي جداً
0.79 - 0.60	عالي
0.59 - 0.40	متوسط
0.39 - 0.20	ضعيف
0.19 - 0.01	ضعيف جداً
0	لا توجد علاقة

الجدول (9-10) القيم الجدولية لدلالة معامل الارتباط

درجة الحرية ن - 2	مستوى الدلالة		درجات الحرية ن - 2	مستوى الدلالة	
	0.05	0.01		0.05	0.01
1	0.997	1.000	24	0.388	0.496
2	0.950	0.990	25	0.381	0.487
3	0.878	0.959	26	0.374	0.478
4	0.811	0.917	27	0.367	0.470
5	0.754	0.874	28	0.361	0.463
6	0.707	0.834	29	0.355	0.456
7	0.666	0.791	30	0.349	0.449
8	0.632	0.765	35	0.325	0.418
9	0.602	0.735	40	0.304	0.393
10	0.576	0.707	45	0.288	0.372
11	0.553	0.684	50	0.273	0.354
12	0.532	0.661	60	0.250	0.325
13	0.514	0.641	70	0.233	0.302
14	0.497	0.623	80	0.217	0.283
15	0.482	0.606	90	0.205	0.267
16	0.468	0.590	100	0.195	0.254
17	0.456	0.575	125	0.174	0.228
18	0.444	0.561	150	0.159	0.208
19	0.433	0.549	200	0.138	0.181
20	0.423	0.537	300	0.113	0.148
21	0.413	0.526	400	0.098	0.128
22	0.404	0.515	500	0.088	0.115
23	0.396	0.505	1000	0.062	0.081

مراجع الفصل التاسع

1- أبو صالح، محمد صبحي، وآخرين (1995). القياس والتقويم. صنعاء: وزارة التربية والتعليم.
2- أحمد، غريب سيد، جبلي، علي عبد الرزاق؛ محمد، حسن محمد حسن (1992). **الإحصاء الاجتماعي: مبادىء وتطبيقات**. الإسكندرية: دار المعرفة الجامعية.
3- الجادري، عدنان حسين (2003). **الإحصاء الوصفي في العلوم التربوية**. عمان: دار المسيرة.
4- زيتون، عايش (1996). **أساسيات الإحصاء الوصفي**. عمان: دار المسيرة.
5- عدس، عبد الرحمن (1999). **الإحصاء في التربية**. عمان: دار الفكر للطباعة والنشر والتوزيع.
6- علام، صلاح الدين (2000). **تحليل البيانات في البحوث النفسية والتربوية والاجتماعية**. القاهرة: دار الفكر العربي.
7- عودة، أحمد، والخليلي، خليل (1988). **الإحصاء للباحث في التربية والعلوم الإنسانية**. عمان: دار الفكر للنشر والتوزيع.
8- عودة، أحمد؛ والقاضي، منصور (2002). **الإحصاء الوصفي والاستدلالي: الإحصاء الوصفي**. الكويت: مكتبة الفلاح للنشر والتوزيع.
9- محمد، حسن محمد (1992). **أساسيات الإحصاء وتطبيقاته**. الإسكندرية: دار المعرفة الجامعية.
10- المنيزل، عبد الله فلاح (2000). **الإحصاء الاستدلالي وتطبيقاته في الحاسوب باستخدام الرزم الإحصائية SPSS** . عمان: دار وائل للنشر.
11- النبهان، موسى (2005). **أساسيات الإحصاء في التربية والعلوم الإنسانية والاجتماعية**. الكويت: مكتبة الفلاح للنشر والتوزيع.
12- Aron, A., & Aron, E. N. (1999). **Statistics for Psychology (2nd ed).** New Jersy: Prentic Hall, Inc.

13- Cook, D., Grawen, A., & Clarke, G (1999) **Basic Statistical Computing**. London: Edward Arnold Publishers.

14. Gay, L. R., & Airaisan, P. (2000) **Educational Research: Competencies for Aalysis and Application (6th ed).** New Jersy: Prentic Hall, Inc.

15. Glass, G., & Hopking, K. (1996). **Statistical Methods in Education and Psychology.** New Jersy: Prentice Hall, Inc.

16. Howell, D. C. (1992) **Statistical Methods for Psychology**. Belmont, California: Duxbury Press.

17. Kirk, R. E (1990). **Statistics: An Introduction**. Fort Work; Holt, Rinehart and Winston.

18. Misanin, J. R., & Hinderliter, C. F. (1991). **Fundamentals of Statistics for Psychology Students**. New York: harpercollins Publishers Inc.

19. Vaughan, E. D. (1998). **Statistics: Tools For Understanding Data in the Behavioral Sceinces**. New Jersy: Prentic - Hall, Ine.

الفصل العاشر

البرمجيات المستخدمة في القياس والتقويم

أولاً- برنامج SPSS

ثانياً- برنامج ITEMAN

ثالثاً- برنامج Hot Potatoes

10

10

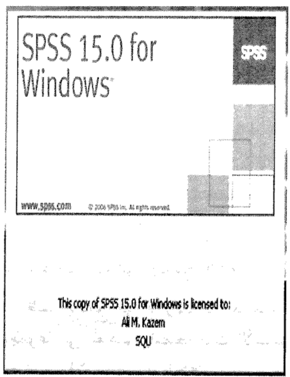
الفصل العاشر

البرمجيات المستخدمة في القياس والتقويم

مقدمة

لقد شاع استخدام الحاسوب في مختلف جوانب ميادين المعرفة، والقياس والتقويم هو أحد المجالات التي ساهم الحاسوب في تطوره كثيرا؛ فظهرت العديد من البرمجيات الحاسوبية في شتى مجالات القياس والتقويم وتطبيقاته المختلفة. وفي هذا الفصل سوف يقتصر الحديث على تطبيقات الرزمة الاحصائية في العلوم الاجتماعية SPSS؛ وبرنامج تحليل المفردة ITEMAN، فضلاً عن برنامج خاص بحوسبة الاختبارات الورقية وتحويلها إلى الكترونية اسمه Hot Potatose.

أولاً - برنامج SPSS



ترجع بدايات ظهور البرنامج الإحصائي SPSS (الرزمة أو الحقيبة الإحصائية في العلوم الاجتماعية) إلى عام 1975، عندما صدرت أول نسخة منه وكانت تعمل على أجهزة الحاسوب المركزية (Main-frame)، بعدها صدرت نسخة مطورة اسمها SPSS-X عام 1983 تعمل بنفس بيئة عمل الإصدار السابق، ولكنها تميزت بسهولة الاستخدام وسرعة الإنجاز.

وفي عام 1990 صدرت أول نسخة من البرنامج تعمل على الحاسب الشخصي PC، وكان اسمها SPSS/PC90+، وفي عام 1993 صدرت أول

نسخة تعمل على نظام النوافذ Windows، وتوالت بعد ذلك الإصدارات بشكل سنوي تقريبا، وأحدث نسخة متاحة الآن هي: SPSS Version15.0، ويمكن القول أن الإصدارات من (10-15) تتشابه إلى حد كبير من حيث طريقة الاستخدام، ونجد أن خبرة المُستخدم في أحد هذه الإصدارات الستة، تجعله قادر على التعامل بفعالية مع الإصدارات الأخرى.

ومما تجدر الإشارة إليه، أن الإصدار الرابع عشر والخامس عشر من هذا البرنامج يتميزان عن الاصدارات السابقة بميزات عديدة، لعل من أبرزها إمكانية فتح أكثر من ملف في وقت واحد؛ حيث يكون لكل ملف نافذة مستقلة خاصة به، في حين لم يكن ذلك ممكنا في جميع الإصدارات السابقة، وإعلام المستخدم في واجهة النتائج بالأوامر التي أعطيت لتنفيذ مهام جديدة، فضلا عن ميزات أخرى تتعلق بسرعة إجراء التحليلات الإحصائية، وطريقة إجراء الأشكال البيانية.

وبشكل عام يمكن القول أن برنامج SPSS يمثل بيئة إحصائية شاملة، وسهل الاستخدام، ويتسم بالمرونة، حيث يسهل استخدامه في إدارة البيانات ابتداء من إدخالها وحفظها واستعادتها وتحليلها وصولا الى استخراج النتائج والمساعدة في تفسيرها، ويستطيع البرنامج قراءة البيانات من قواعد البيانات Databases أو أنظمة جداول البيانات Spread Sheets وبصيغ متعددة، ويمكن إجراء تحليلات إحصائية متقدمة (معقدة) بأمر واحد بسيط في وقت قصير جدا. كما يمكن الحصول على إحصاءات وصفية Descriptive Statistics، ورسومات بيانية (أشكال توضيحية) Charts، وجداول تقاطعية Cross tabulation متقدمة، فضلاً عن الإحصاءات الاستدلالية المختلفة .Inferential Statistics

مفاهيم أساسية في برنامج SPSS

طبقا لبرنامج SPSS وكذلك العديد من البرامج الإحصائية الأخرى، هناك مجموعة من المفاهيم والمصطلحات الأساسية:

1- يسمى الأشخاص أو الطلاب الذين يجيبون عن اختبار تحصيلي أو استبيان معين، حالة Case.

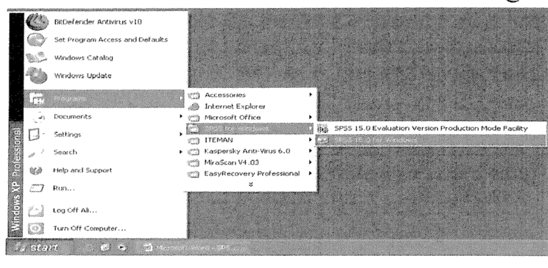
2- يسمى كل سؤال أو فقرة في الاختبار متغير Variable.

3- تسمى إجابات الأشخاص على الأسئلة بقيم المتغيرات Values of variables.

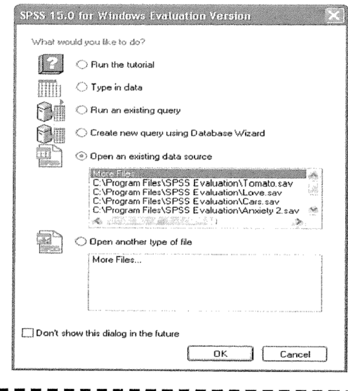
4- تسمى نقطة تقاطع العمود مع الصف بالخلية Cell وكل خلية تمثل قيمة متغير معين لحالة معينة.

فتح برنامج SPSS

أنقر على قائمة البدء Start في جهاز الحاسوب، واختر من قائمة البرامج برنامج SPSS 15.0 for Windows.



ستفتح واجهة البرنامج وفي داخلها نافذة تتضمن السؤال الآتي: (What would you like to do?) ماذا تريد أن تفعل؟).



يتضمن صندوق الحوار السابق ستة إجابات (أو اختيارات) هي:

1	Run the tutorial	جولة تعريفية داخل البرنامج.
2	Type in data	فتح ملف بيانات فارغ.
3	Run an existing query	استعلام عن بيانات موجودة في البرنامج.
4	Create new query using Database Wizard	استيراد بيانات من برامج أخرى غير برنامج SPSS، مثل: dBASE, Excel, MS Access Database
5	Open an existing data source	فتح آخر ملفات بيانات تم التعامل معها في البرنامج، وذلك من خلال اختيار الملف والنقر على .OK
6	Open another type of file	فتح آخر ملفات خاصة بالنتائج تم التعامل معها في البرنامج، وذلك من خلال اختيار الملف والنقر على OK

وهناك خيار على الجانب الأيسر في أسفل صندوق الحوار Don't show this dialogue in the future، يتيح هذا الخيار عدم فتح هذا الصندوق الحواري عند فتح البرنامج من جديد، وسيؤدي إلى فتح واجهة محرر البيانات مباشرة SPSS Editor View

جولة في برنامج SPSS

يحتوي برنامج SPSS على أربعة واجهات (أو شاشات)، لكل واجهة وظائف معينة، يستطيع المستخدم من خلالها التعامل مع تلك الوظائف عن طريق مجموعة من القوائم Menus، والأيقونات Icons وفيما يأتي عرض لهذه الواجهات، يعقبه عرض لمهام ووظائف القوائم والأيقونات.

الأول - واجهات برنامج SPSS

هناك واجهات عديدة في برنامج SPSS، ليس من أهداف هذا الكتاب التعرض لها، ولكن بقدر تعلق الأمر باستخدامات البرنامج في مجال القياس والتقويم لابد من التعرض لأهمها، وهي:

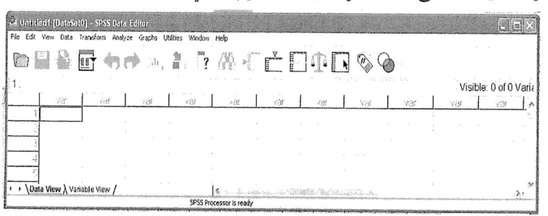
(1) الواجهة الرئيسة Main View

تضم الواجهة الرئيسة للبرنامج، واجهتين فرعيتين، الأولى تفتح تلقائيا عند فتح البرنامج، وتسمى واجهة محرر البيانات Data View، والواجهة الثانية تسمى Variable View يتم الانتقال اليها عن طريق النقر على أيقونتها، وفيما يأتي عرض لكل واجهة.

ile Edit Configure Analyze Help
Data File
23:28:54

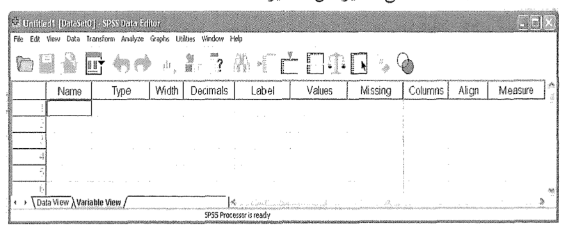
أ- واجهة محرر البيانات Data Editor View

تعد واجهة محرر البيانات الصفحة الرئيسة للبرنامج؛ حيث أنها تفتح مباشرة عند فتح البرنامج، ومن خلالها يستطيع المستخدم استدعاء ملف سابق (مهما كان نوعه)، أو عمل ملف جديد. وواجهة محرر البيانات هي صفحة تتكون من مصفوفة من الخلايا (أعمدة وصفوف). أعمدة الخلايا معنونة بـ "المتغير" Var، والصفوف لها أرقام متسلسلة. وإما قدرة البرنامج على التعامل مع المتغيرات والحالات التي يمكن إدخالها إليه؛ فهي غير محددة، اي أنه يمكن إدخال عدد غير نهائي من الحالات، وعدد غير نهائي من المتغيرات، وذلك يعتمد على سعة ذاكرة جهاز الحاسوب الذي تستخدمه.



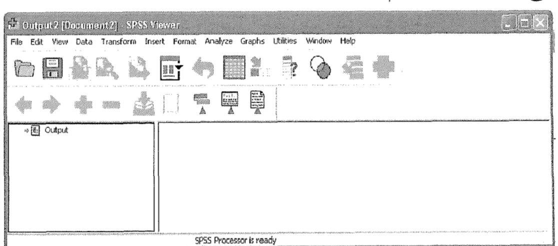
ب- واجهة المتغيرات **Variable View**

تعد واجهة المتغيرات الصفحة الرئيسة الثانية من حيث الأهمية بعد واجهة البيانات، ويمكن للمُستخدم الوصول إليها عن طريق النقر على Variable View الموجودة أسفل واجهة البيانات على الجهة اليسرى. كما يمكن العودة الى واجهة البيانات عن طريق النقر على Data View الموجودة الى جانبها. وهذه الواجهة مهمة عند إنشاء ملف جديد، وفيها معلومات قاموسية Dic-tionary Information لكل متغير من المتغيرات.



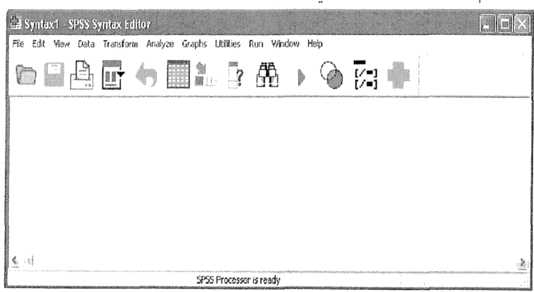
(2) واجهة المخرجات **Out Put View**

يتم في هذه الواجهة عرض نتائج التحليلات الإحصائية التي تم تنفيذها، وفيها يستطيع المستخدم النقر المزدوج على أي نتيجة (جدول، أو شكل بياني) لتفتح له خيارات التحكم فيه من حيث التعديل والإضافة.



(3) واجهة محرر التعليمات والأوامر Syntax View

وهي الواجهة التي يتم من خلالها كتابة التعليمات أو الخطوات للإجراءات الإحصائية التي يقوم بتنفيذها البرنامج (تقابل المعادلات الإحصائية)، ونظرا لتعقد هذه الإجراءات، فإنها جاهزة في قائمة التحليلات Analyze، أو غيرها من شريط الأدوات، ومن الناحية العملية لا يحتاج المستخدم الى هذه النافذة إلا في بعض الحالات.



الثاني - قوائم واجهات برنامج SPSS

أ- قوائم واجهة البيانات والمتغيرات Data- Variables View Menus

File	Edit	View	Data	Transform	Analyze	Graphs	Utilities	Window	Help

(1) قائمة ملف File Menu

وتتضمن آليات التعامل مع الملفات من حيث: عمل ملف جديد، أو فتح ملف مُخزّن سابقا، أو تخزين الملف الحالي، أو طباعته، كذلك استيراد ملف بيانات من إصدارات SPSS المختلفة، أو من نظام إحصائي آخر، كنظام: (Systat, Excel, Lotus, SYLK, dBASE, SAS, Text, Data)؛ أو تصدير الملف

الحالي إلى إصدارات SPSS المختلفة أو الى نظام إحصائي آخر مثل: (Tab- de- limited, Fixed ASCII, Excel, 1-2-3 Rel, SYLK, dBASE, SAS)، فضلا عن الخروج من النظام بشكل كلي.

File Edit View Data Transform Analyz
New
Open
Open Database
Read Text Data...
Close Ctrl+F4
Save Ctrl+S
Save As...
Save All Data
Export to Database...
Mark File Read Only
Rename Dataset...
Display Data File Information
Cache Data...
Stop Processor Ctrl+.
Switch Server...
Print Preview
Print... Ctrl+P
Recently Used Data
Recently Used Files
Exit

(2) قائمة تحرير **Edit Menu**

وفيها العديد من الأوامر مثل: التراجع، التقدم، القص، النسخ، اللصق، الحذف، البحث عن قيمة معينة، والخيارات المتعلقة بإعداد النظام.

Edit View Data Transform
Undo Ctrl+Z
Redo Ctrl+R
Cut Ctrl+X
Copy Ctrl+C
Paste Ctrl+V
Paste Variables...
Clear Del
Insert Variable
Insert Cases
Find... Ctrl+F
Go to Case...
Options...

(3) قائمة عرض View Menu

وفيها يمكن التحكم في شريط الأدوات، والتحكم في خطوط شبكة واجهة البيانات أو واجهة المتغيرات (إخفائها وإظهارها)، والتحكم في الخط (نوعه وحجمه ونمطه)، وكذلك في إظهار وإخفاء دلالات القيم التي أعطيت للمتغيرات.

View Data Transform Analyze Graphs
✓ Status Bar
Toolbars...
Fonts...
✓ Grid Lines
Value Labels
Display Custom Attributes...
Variables Ctrl+T

(4) قائمة البيانات Data Menu

من خلال الأوامر الموجودة فيها يتم تعريف البيانات، وإضافة متغير أو حالة إلى البيانات، وترتيب الحالات تنازليا أو تصاعديا، ودمج الملف مع ملف آخر، اوقسمة الملف حسب متغير معين إلى ملفين أو أكثر، وكذلك إعطاء أوزان للمتغيرات.

Data Transform Analyze Graphs
Define Variable Properties...
Copy Data Properties...
New Custom Attribute...
Define Dates...
Define Multiple Response Sets...
Validation
Identify Duplicate Cases...
Identify Unusual Cases...
Sort Cases...
Transpose...
Restructure...
Merge Files
Aggregate...
Orthogonal Design
Copy Dataset
Split File...
Select Cases...
Weight Cases...

(5) قائمة التحويل (أو التحوير والتغيير) Transform Menu

تحتوي هذه القائمة على العديد من الأوامر التي يتم من خلالها تحويل البيانات من حالة إلى أخرى، بإجراء العمليات الحسابية المختلفة وذلك باستخدام الدوال الإحصائية الجاهزة، وإعادة ترميز البيانات في المتغير نفسه أو في متغير جديد، وإعطاء فئات للمتغيرات، وتحديد الرتب للحالات، وغيرها.

Transform Analyze Graphs Utilities Window

- Compute Variable...
- Count Values within Cases...
- Recode into Same Variables...
- Recode into Different Variables...
- Automatic Recode...
- Visual Binning...
- Optimal Binning...
- Rank Cases...
- Date and Time Wizard...
- Create Time Series...
- Replace Missing Values...
- Random Number Generators...
- Run Pending Transforms Ctrl+G

(6) قائمة التحليلات الاحصائية Analyze Menu

تعد الأوامر الموجودة في هذه القائمة العمود الفقري او (العقل المدبر Master Mind) للبرنامج، حيث تتضمن جميع الوسائل (أو الاختبارات) الإحصائية التي يمكن ان يتضمنها أي كتاب في الإحصاء تقريباً، سواء كانت إحصاءات بسيطة simple مثل: التكرارات ومقاييس النزعة المركزية ومقاييس التشتت والجداول بأنواعها . . . . . الخ، أو كانت تلك الإحصاءات متقدمة advanced مثل: تحليل التباين والانحدار والتحليل العاملي . . . . الخ.

Analyze Graphs Utilities Win

- Reports
- Descriptive Statistics
- Tables
- Compare Means
- General Linear Model
- Generalized Linear Models
- Mixed Models
- Correlate
- Regression
- Loglinear
- Classify
- Data Reduction
- Scale
- Nonparametric Tests
- Time Series
- Survival
- Multiple Response
- Missing Value Analysis...
- Complex Samples
- Quality Control
- ROC Curve...

(7) قائمة الرسوم البيانية (الأشكال) Graphs Menu

وتتضمن مختلف الرسوم البيانية المستخدمة في العرض البياني مثل : الأعمدة البيانية والخط البياني والدائرة البيانية ومخطط الانتشار والسلاسل الزمنية وغيرها كثير.

Graphs Utilities Win
Chart Builder...
Interactive
Legacy Dialogs
Map

(8) قائمة الخدمات (أو الادوات) Utilities Minu

وتتضمن معلومات تفصيلية عن ملف البيانات الحالي بما يحويه من متغيرات، وكذلك وضع أو تجميع تلك المتغيرات في مجموعات على أساس معيار معين، فضلا عن المُحرر الرئيسي للبيانات.

Utilities Window Help
Variables...
OMS Control Panel...
OMS Identifiers...
Data File Comments...
Define Variable Sets...
Use Variable Sets...
Run Script...
Menu Editor...

(9) قائمة النافذة Window Menu

وفيها يستطيع المُستخدم التنقل بين النوافذ (أو الشاشات) المفتوحة في النظام، وهي ثلاث نوافذ (البيانات والمتغيرات Data & Variables، والمخرجاتOutput، ومحرر التعليمات Syntax) فمثلا يمكن الانتقال من نافذة النتائج او محرر البيانات في حال استخدامها، فضلا عن إمكانية التحكم في حجم تلك النوافذ.

Window Help
Split
Minimize All Windows
✔ 1 Untitled1 [DataSet0] - SPSS Data Editor

(10) قائمة المساعدة Help

مهما تمكن المُستخدم من برنامج SPSS أو أي برنامج آخر، إلا انه يواجه بعض المواقف التي يحتاج فيها إلى توجيه أو استشارة، والمساعدة هنا تقدم للمُستخدم كشف مفصل بما يتضمنه النظام من وسائل إحصائية، وكيفية استخدام كل وسيلة خطوة خطوة، وباختصار المساعدة تمثل دليل النظام الذي لا يمكن لاي فرد الاستغناء عنه.

Help
Topics
Tutorial
Case Studies
Syntax Guide
Statistics Coach
Command Syntax Reference
SPSS Developer Central
Algorithms
SPSS Home Page
About...
License Authorization Wizard
Register Product

ب- قوائم واجهة المخرجات Out Put View Menus

Output2 [Document2] - SPSS Viewer
File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help

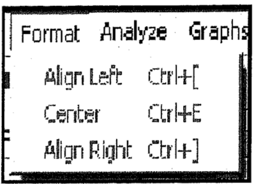
تتشابه قوائم واجهة البيانات وواجهة المتغيرات، مع قوائم واجهة المخرجات، ولكن هناك قائمتين جديدتين في واجهة المخرجات هما القائمة رقم 6 و7 ، وفيما يلي عرض لهما:

(6) قائمة إدراج Insert Menu

في هذه القائمة يمكن إدراج عنوان للنتيجة أو عنوان للصفحة، أو إضافة نص وغير ذلك من الإضافات سواء باللغة العربية أو الانجليزية.

Insert Format Analyze
Page Break
Clear Page Break
New Heading
New Title
New Page Title
New Text
Interactive 2-D Graph
Interactive 3-D Graph
New Map
Text File...
Object...

(7) قائمة تنسيق Format Menu



تتضمن هذه القائمة محاذاة النص والجدول في جهة اليمين واليسار والوسط، ويتم ذلك بعد وضع مؤشر على النتيجة وتحديدها، ومن ثم التحكم فيها.

ج - قوائم واجهة محرر التعليمات Syntax View Menu

File Edit View Data Transform Analyze Graphs Utilities Run Add-ons Window Help

تتشابه قوائم واجهة محرر البيانات مع قوائم واجهة المخرجات وواجهة البيانات والمتغيرات، ولكن واجهة محرر التعليمات تحتوي على قائمة جديدة هي القائمة رقم 9، وفيما يلي عرض لها:

(9) قائمة نفّذ Run Menu

Run Add-ons Window
All
Selection
Current Ctrl+R
To End

وهي قائمة تتفرد بها واجهة محرر التعليمات، يتم من خلالها تنفيذ الأوامر التي كتبها او اعدها المستخدم كبديل لاستخدام القوائم الجاهزة.

الثالث - أيقونات واجهات برنامج SPSS

هناك مجموعة من الأيقونات المتشابهة في جميع واجهات البرنامج الأربع، وهناك بعض الأيقونات المتخصصة بكل واجهة، وجميع هذه الأيقونات تساعد المستخدم على سرعة انجاز العمل، وتقلل من اعتماده على القوائم. وفيما يأتي عرض لهذه الأيقونات.

(1) أيقونات واجهة البيانات والمتغيرات

تحوي واجهة البيانات وواجهة المتغيرات على عدد من الإيقونات المتشابهة، يبلغ عددها (17) أيقونة، بعضها نشط (Active لونها واضح)، وعددها (9) أيقونة، وبعضها غير نشط (Inactive لونها خافت، غير واضح) وعددها (8) أيقونات يتم تنشيطه أثناء التعامل مع البيانات.

	Open File	فتح ملف موجود سابقا في مكان ما.
	Save File	حفظ ملف.
	Print	طباعة ملف.
	Dialog Recall	أظهار آخر 12 عملية تحليل إحصائي.
	Undo	تراجع عن آخر تعديل.
	Redo	تراجع عن تراجع.
	Go To Chart	الذهاب إلى الرسوم البيانية والأشكال.
	Go To Case	الذهاب إلى حالة معينة.
	Variables	تفاصيل المتغيرات.
	Find	بحث عن حالة ضمن متغير معين.
	Insert Case	إدراج حالة جديدة في مكان معين
	Insert Variable	إدراج متغير جديد في مكان معين.
	Split File	تجزئة ملف.
	Weight Case	إعطاء أوزان لحالات معينة.
	Select Case	اختيار مجموعة من الحالات.
	Value Lables	إظهار دلالات قيم المتغيرات (إذا تم تحديدها).
	Use Sets	استخدام مجموعة جزئية من المتغيرات.

(2) أيقونات واجهة المخرجات (النتائج)

تضم واجهة المخرجات (22) أيقونة، (10) منها نشطة، و(12) غير نشطة، وهي:

	Open File	فتح ملف موجود سابقا في مكان ما
	Save File	حفظ ملف.
	Print	طباعة ملف.
	Preview Print	معاينة قبل الطباعة.
	Export	تصدير تفاصيل النتائج بلغة HTML مثلا مع تفاصيل عن محرر البيانات.
	Dialog Recall	إظهار آخر 12 عملية تحليل إحصائي.
	Undo	تراجع عن آخر تعديل.
	Go To Data	الذهاب إلى البيانات.
	Go To Case	الذهاب إلى خالة معينة.
	Variables	تفاصل المتغيرات.
	Use Variables Sets	مجموعة المتغيرات المستخدمة.
	Select Last Out Put	اختيار آخر نتيجة تم حسابها.
	Designate Window	النافذة المعينة (المرشحة).
	Promote	إلى الامام.
	Demote	إلى الخلف.
	Expand	توسيع.
	Collapse	تقليل.
	Show	عرض.
	Hide	إخفاء.
	Insert Heading	إدراج رؤوس موضوعات.
	Insert Title	إدراج عنوان.
	Insert Text	إدراج نص.

(3) أيقونات واجهة محرر التعليمات.

تحوي واجهة محرر البيانات (13) أيقونة، (10) منها نشطة، و(3) غير نشطة، وهي:

	Open File	فتح ملف موجود سابقا في مكان ما.
	Save File	حفظ ملف.
	Print	طباعة ملف.
	Dialog Recall	إظهار آخر 12 عملية تحليل إحصائي.
	Undo	تراجع عن آخر تعديل.
	Go To Data	الذهاب إلى البيانات.
	Go To Case	الذهاب الى حالة معينة.
	Variables	تفاصيل المتغيرات.
	Finding	بحث واستبدال.
	Run Selection	تنفيذ أوامر مختارة.
	Use Variable Sets	مجموعات المتغيرات المستخدمة.
	Syntax Help	مساعدة خاصة بمحرر التعليمات.
	Designate Window	النافذة المعينة (المرشحة).

ادخال البيانات في برنامج SPSS

1 - فحص الإجابات (تدقيق إجابات الطلبة)

قبل الشروع في عملية إدخال البيانات في البرنامج، يتعين فحص إجابات الطلاب على الامتحان، وذلك للتأكد من عدم وجود نمط معين للإجابة، وأن الطلاب قد أجابوا على الاختبار بصورة صحيحة، ولا توجد اكثر من إجابة واحدة لكل سؤال.

إن البرنامج يتعامل مع البيانات المدخلة بطريقة صحيحة أو خاطئة سواء، لذا نجد من المهم التأكد من إدخال البيانات بالشكل الصحيح، كما يتعين تطبيق قاعدة محددة في التعامل مع الحالات التي أجاب فيها الطالب أكثر من إجابة، كأن تترك هذه الأسئلة بدون إجابة وتعتبرها قيم مفقودة Missing، أو تعتبر الطالب الذي ترك الأسئلة بأنه غير متيقن من إجابته وتعاملها معاملة الإجابات الخاطئة.

2 - ترميز المتغيرات

قبل إدخال البيانات كملف في برنامج SPSS، يتعين إعطاء رموز لأسئلة الاختبار ومتغيراته الديمغرافية كجنس الطالب وصفه ورقمه وغير ذلك، وتسمى هذه العملية بالترميز Coding.

يمكن لبرنامج SPSS أن يتعامل مع الإرقام والحروف في آن واحد، فلو أخذنا متغير الجنس مثلا فيمكننا أن نشير إلى الذكور بالرقم 1 والإناث بالرقم 2، أو نشير إلى الذكور بـ M والإناث بـ F، ولكن يفضل دائماً استخدام الإرقام Numeric بدلا من الحروف String؛ لسهولة تعامل البرنامج معها، ولأن إدخال الرموز يحتاج إلى القيام ببعض الخطوات الإضافية التي لا تحتاجها الأرقام. وبما أن الاختبار يحتوي على عدة أنواع من الأسئلة، فإن عملية الترميز تأخذ طابعا مختلفا باختلاف نوع الأسئلة، وفيما يأتي أمثلة لبعض أنواع الأسئلة، وكيفية ترميزها:

- الأسئلة التي تسمح باختيار إجابة واحدة (الموضوعية):

من أمثلة هذا النوع أسئلة الصواب والخطأ، وفي هذا النوع من الاسئلة هناك احتمالين للإجابة، فإما أن تكون صحيحة ويرمز لها بالرمز (1) مثلا، أو أن تكون خاطئة ويرمز لها بالرمز صفر (0)، وتجدر الإشارة إلى ان الرمز يمكن ان ياخذ أي قيمة، ولكن من المهم أن يكون الترميز متسقاً (موحداً)، أي استخدام نفس الرمز للإجابة الصحيحة مع جميع الأسئلة ,وكذلك الحال بالنسبة لترميز الأسئلة الخاطئة.

ومن أمثلة هذا النوع أيضاً اسئلة الاختيار من متعدد وذلك عن طريق إعطاء رمز او رقم إجابة الطالب، مثال:

- أي من الكواكب الآتية الاقرب إلى الأرض؟

أ) الزهرة . ب) عطارد . جـ) المريخ. د) زحل.

- الأسئلة المفتوحة (الإنشائية):

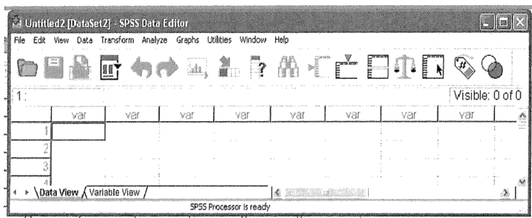
في كثير من الأحيان تمثل الاسئلة المفتوحة جزءاً لا بأس به من الاختبار، وعادة ما نقوم بتصحيح الاسئلة المفتوحة يدويا ومن ثم يتم إدخال درجة الطالب في السؤال إلى الحاسب الآلي.

3 - إدخال البيانات في برنامج SPSS

هناك طريقتان لإدخال البيانات في برنامج SPSS، ولهذا الغرض سوف نستخدم إجابات عشرون طالباً وطالبة في امتحان مكون من خمسة عشر سؤالا من نوع الاختيار من متعدد والموضحة في أدناه:

أسماء الطلبة	أرقام الإجابة وإجابات الطلبة														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1 مازن	b	b	d	b	c	d	a	a	c	b	c	d	c	a	a
2 محمد	c	a	b	b	d	b	a	a	c	b	b	d	d	a	b
3 أحمد	b	a	d	b	c	d	a	c	c	b	c	d	b	a	a
4 فاطمة	b	c	d	b	c	d	b	b	c	b	c	d	d	b	c
5 لبنى	b	c	d	b	c	d	b	b	c	b	c	d	d	a	d
6 هدى	b	a	b	b	c	b	b	a	c	c	a	d	a	a	a
7 قيس	b	a	d	b	d	d	a	d	b	b	c	d	d	b	a
8 أماني	b	a	b	b	c	b	b	a	c	b	b	d	d	b	a
9 سالم	b	a	d	b	c	b	b	a	a	b	d	d	d	a	a
10 علي	b	a	c	b	d	b	b	a	c	c	c	d	d	a	a
11 إيمان	b	b	d	b	a	b	b	b	a	b	b	d	a	b	b
12 وداد	b	a	c	b	c	b	b	d	b	d	d	c	b	b	c
13 ليلى	b	a	d	c	c	b	b	a	d	b	b	c	d	c	a
14 وليد	c	c	a	b	d	c	c	a	c	b	b	a	a	a	b
15 ماهر	b	a	b	b	b	b	b	a	a	b	b	d	a	a	c
16 سلطان	b	b	d	a	d	b	b	a	a	b	b	d	c	a	d
17 عبير	a	c	c	b	c	a	a	a	d	b	b	d	c	a	d
18 عنود	b	b	d	c	b	b	b	d	c	b	b	c	b	d	a
19 وعد	b	a	a	b	c	b	b	a	c	b	b	d	d	a	a
20 عزيز	b	c	d	d	d	c	b	c	c	a	b	a	a	d	b
مفتاح الإجابة	b	a	d	b	c	d	a	a	c	b	c	d	d	c	a

تعتمد الطريقة الأولى على إدخال أسماء المتغيرات والبيانات (رقم السؤال وإجابة الطالب في هذه الحالة) في صفحة المحرر مباشرة. فبعد أن تفتح صفحة المحرر، تلاحظ في أسفل الصفحة زرين أحدهما معنون بـ Data View والثاني معنون بـ Variable View.

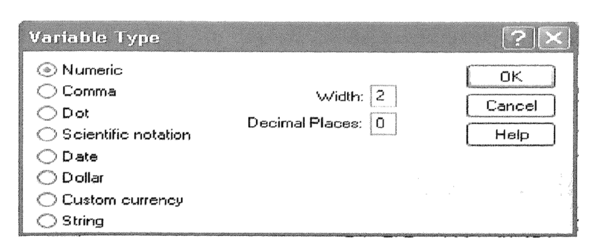


لتعريف المتغيرات او أسئلة الاختبار أنقر على Variable View، يمكن إعطاء أسماء مختصرة للمتغيرات طالما ان كل سؤال له اسم مختلف، وبما أننا نتعامل مع أسئلة امتحان واسماء طلاب، فإننا سوف نستخدم رقم الطالب كتعريف به في العمود المعنون (ID)، ونرمز للذكور بالرقم 1 وللإناث بالرقم 2 في العمود المعنون (Sex)، كما سنستخدم الرمز q متبوعا بأرقام متسلسلة من 1 إلى ن (ن= عدد أسئلة الامتحان).

لكل متغير نقوم بتعريفه عشرة خصائص تسمى المعلومات القاموسية Dic-tionary Infomation، الخاصية الأساس هي Name، ويكفي تعرفها والانتقال بالسهم للمتغير الثاني؛ حيث أن البرنامج تلقائياً يقوم بتحديد الخصائص الأخرى، ولكن هناك بعض الخصائص لابد من تحديدها مثل رمز الذكور ورمز الإناث، وغيرها. وفيما يلي تعريف بكل خاصية:

1- Name وهو الاسم المختصر للمتغير Short Name، ويشترط أن يبدأ بحرف.

2- Type نوع المتغير، تلقائياً البرنامج يعتبره رقمي Numeric، ويمكن أن يكون حرفي String، أو أي نوع آخر (هناك ثمانية أنواع للمتغيرات).



3- Width عدد الأرقام الصحيحة للمتغير، وتلقائيا البرنامج يعطي العدد 8 ولكن يمكن تعديل حسب الحاجة.

4-Decimals عدد الارقام العشرية في المتغير، وتلقائيا البرنامج يعطي رقمين عشريين، ويمكن تعديله حسب حالة المتغير.

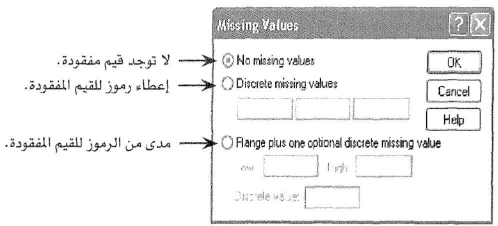
5- Label الاسم الطويل للمتغير Long Name، وفيه نستطيع ان نكتب سؤال الاختبار، او فقرة الاستبيان، علما ان كتابة الاسم الطويل للمتغير سيؤدي إلى إظهاره في جميع المعالجات والتحليلات الإحصائية، وفي حالة عدم كتابته سيظهر الاسم القصير في جميع النتائج.

6- Values وهي وصف للقيم المرمزة، في حالة متغير النوع، أعطينا الذكور 1، والإناث 2، ننقر على الخلية المقابلة لمتغير النوع سيفتح صندوق الحوار، نكتب فيه المعلومات الآتية:

Value Labels

Value Labels
Value:
Label:
Add
Change
Remove
1 = "male"
2 = "female"
OK
Cancel
Help

7- Missing القيم المفقودة، النقر على هذه الخلية يؤدي إلى فتح صندوق حوار فيه ثلاثة خيارات:



8-Columns أتساع عمود بيانات المتغير.

9- Align محاذاة العمود، وفيه ثلاثة خيارات: اليسار، واليمين، والوسط.

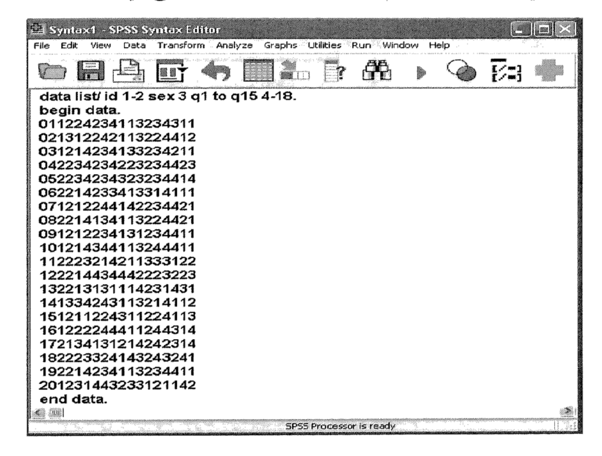
10- Measure مستوى القياس، وفيه ثلاثة خيارات: فئوي Scale، ورتبي Ordinal وأسمي Nominal.

بعد الانتهاء من تعريف خصائص المتغير نبدأ بإدخال البيانات، وعليه يتعين الانتقال الى واجهة البيانات بالنقر على (Data View)، ووضع مؤشر الفأرة على الخلية الأولى من العمود الأول والسطر الأول. بحيث ندخل إجابات كل طالب مرة واحدة، سؤال بعد آخر، وذلك بأن نعطي الرقم 1 للأجابة a و2 للإجابة c، و4 للإجابة d، وتحريك المؤشر الى اليمين بواسطة السهم في لائحة المفاتيح للانتقال من خلية لأخرى. ويلاحظ بأن هذه الطريقة سهلة، ولكن يعيبها طول الوقت النسبي الذي تستغرقه. وهذا يأخذنا الى الطريقة الثانية في إدخال البيانات. والتي تعتمد على لغة البرمجة الخاصة بـ SPSS، وفيها يتم كتابة البيانات على هيئة أوامر محدد، وهي أسرع نسبيا من الطريقة الأولى، وتتلخص في الخطوات التالية:

1- أختر Syntax تحت Open New من قائمة File

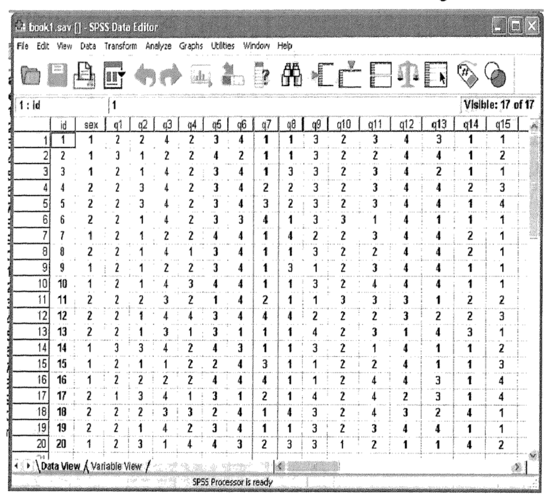
2- في واجهة محرر التعليمات نقوم بتعريف البيانات في السطر الأول،

وذلك بكتابة عبارة Data list متبوعة بأسماء المتغيرات (الأسئلة) بحيث يكون كل سؤال أمامه عدد الخانات أو الأعمدة المطلوبة لأدخال الدرجة (آحاد، أو عشرات، أو مئات)، فإذا كانت الإجابة تمثل رقما واحداً ففي هذه الحالة نحتاج لعمود واحد، وإذا كانت أكثر من ذلك فنحتاج إلى أكثر من عمود وهكذا، وفي السطر التالي نقوم بكتابة عبارة begin data، وفي السطر الذي يليه نقوم بإدخال بيانات كل حالة في سطر مستقل، وفي السطر الأخير نقوم بكتابة end data كما هو موضح في أدناه.



ومما تجدر الإشارة إليه أن ترك المسافة في السطر الأول مهم جدا، كما أن وجود النقطة في نهاية الأوامر مهمة، لكنها غير مطلوبة في نهاية أسطر البيانات، بعد الانتهاء من عملية الإدخال احفظ ملف محرر التعليمات عن طريق النقر على File ثم Save as، ثم حدد اسم الملف ومكانه، بعدها أنقر على OK

بعد ذلك اختر All من القائمة Run في صندوق الحوار للإطلاع على ملف البيانات، أو أنقر على أيقونة ▶ ثم انتقل إلى محرر البيانات ستجد البيانات بالشكل المبين أدناه.



	id	sex	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
1	1	1	2	2	4	2	3	4	1	1	3	2	3	4	3	1	1
2	2	1	3	1	2	2	4	2	1	1	3	2	2	4	4	1	2
3	3	1	2	1	4	2	3	4	1	3	3	2	3	4	2	1	1
4	4	2	2	3	4	2	3	4	2	2	3	2	3	4	4	2	3
5	5	2	2	3	4	2	3	4	3	2	3	2	3	4	4	1	4
6	6	2	2	1	4	2	3	3	4	1	3	3	1	4	1	1	1
7	7	1	2	1	2	2	4	4	1	4	2	2	3	4	4	2	1
8	8	2	2	1	4	1	3	4	1	1	3	2	2	4	4	2	1
9	9	1	2	1	2	2	3	4	1	3	1	2	3	4	4	1	1
10	10	1	2	1	4	3	4	4	1	1	3	2	4	4	4	1	1
11	11	2	2	2	3	2	1	4	2	1	1	3	3	3	1	2	2
12	12	2	2	1	4	4	3	4	4	4	2	2	2	3	2	2	3
13	13	2	2	1	3	1	3	1	1	1	4	2	3	1	4	3	1
14	14	1	3	3	4	2	4	3	1	1	3	2	1	4	1	1	2
15	15	1	2	1	1	2	2	4	3	1	1	2	2	4	1	1	3
16	16	1	2	2	2	2	4	4	4	1	1	2	4	4	3	1	4
17	17	2	1	3	4	1	3	1	2	1	4	2	4	2	3	1	4
18	18	2	2	2	3	3	2	4	1	4	3	2	4	3	2	4	1
19	19	2	2	1	4	2	3	4	1	1	3	2	3	4	4	1	1
20	20	1	2	3	1	4	4	3	2	3	3	1	2	1	1	4	2

ملاحظة مهمة على طريقة إدخال البيانات

إن عملية إدخال البيانات إحدى المهارات المهمة لمستخدم برنامج SPSS وغيره من البرامج الإحصائية، لذلك فإن السرعة والدقة في عملية الإدخال مؤشران مهمان للحكم على مستوى تلك المهارة، لذلك ننصح بالتدريب المستمر، وإدخال البيانات مباشرة من أوراق الامتحان الى ملف البيانات دون تفريغها في ورقة ثم إدخالها.

4 - التأكد من صحة إدخال البيانات :

بعد الانتهاء من عملية إدخال البيانات Data Entry، يتعين مراجعة البيانات المدخلة أو ما يعرف بتنظيف او تهذيب البيانات Data Cleaning للتأكد من أن إدخال البيانات بالصورة الصحيحة، وأنه لم تكن هناك اخطاء أثناء عملية الإدخال، ولعمل ذلك قم بالآتي:

1- أعد إدخال أو مراجعة عينة عشوائية من الإجابات.

2- قم بإجراء بعض العمليات الإحصائية (إحصاء وصفي) للتأكد من أن البيانات التي تم إدخالها تقع في المدى الموجود في الاختبار، ولا يوجد رقم خارجه، مثلا وجود الرقم 33 ضمن إجابات الطلبة على الأسئلة يعني انه غير صحيح، ربما حصل بسبب الضغط مرتين على الرقم 3 أثناء عملية الإدخال، وهكذا مع بقية المتغيرات.

5 - حفظ البيانات :

بعد تعريف المتغيرات وإدخال البيانات في محرر البيانات، يمكن القيام بحفظ هذه البيانات بإتباع الخطوات التالية :

1) من قائمة File أختر Save as، أو انقر على

2) في النافذة التي تظهر أمامك، أكتب اسم الملف في المستطيل المعنون بـ File name وليكن أسمك مثلا أو أي اسم تختاره.

3) اختر المكان المطلوب تخزين الملف عليه (a او b أو c) او غير ذلك.

4) أنقر على الزر **OK**

وبذلك فإن بياناتك قد تم حفظها ويمكنك فتح هذا الملف لتعديله أو حذفه أو إضافة ما تشاء من بيانات إليه.

6 - تصحيح الاختبار

لقد تعمدنا إدخال البيانات الخام، أي قبل تصحيح الاختبار، توفيرا للوقت؛ لأن البرنامج يقوم بذلك بكل سهولة ويسر، ولتصحيح الاختبار قم بعمل نسخة أخرى من الملف لتحتفظ بالنسخة الأصلية لاستخدامات قد تستجد مستقبلاً، ثم قم بعمل الآتي:

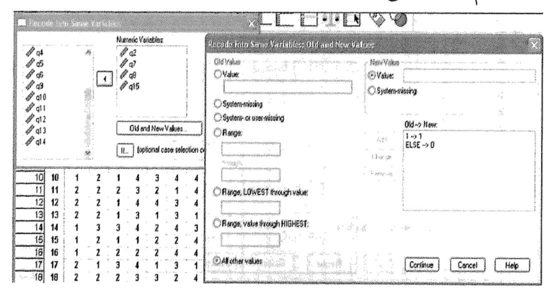
1) بالاستعانة بمفتاح التصحيح، صنف اسئلة الامتحان إلى مجموعات حسب رمز الإجابة الصحيحة (a,b,c,or d) كالآتي:

رمز الاجابة الصحيحة	أرقام الأسئلة
a	15, 8, 7, 2
b	10, 4, 1
c	14, 11, 9, 5
d	13, 12, 6, 3

2) من قائمة تحويل Transform، اختر الأمر Recode Into Same Variables

3) قم بنقل الأسئلة التي كانت إجابتها الصحيحة (1) وهي تقابل الإجابة (a) إلى المربع الموجود على يمين الشاشة، ومن ثم اضغط عل الزر المعنون بـ old and new Values.

4) إدخل الرمز 1 في المربع المعنون Value تحت العنوان New Value، والرقم 1 في المربع المعنون بـ Value تحت العنوان New Value، ثم اضغط على كلمة Add، بعد ذلك انقر على All Other Value تحت العنوان Old Value مع كتابة (0) (صفر) في المربع المعنون بـ Value تحت العنوان New Value، ثم اضغط على كلمة Add



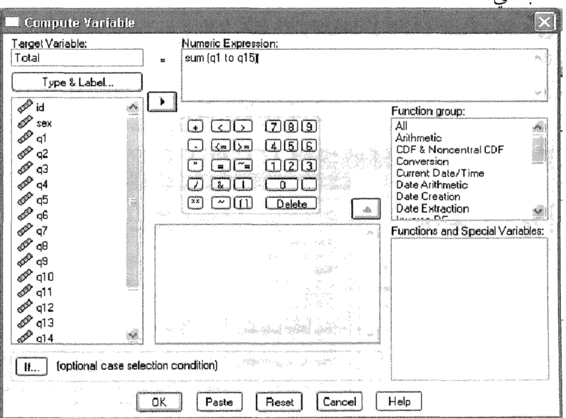
5) اتبع نفس الخطوات السابقة لترميز باقي الإجابات (2،و 3،و4)، وبعد الانتهاء من ذلك ستجد بأن محرر البيانات فيه رقمين، اما صفر ، أو واحد.

7 - التحليل الإحصائي للاختبار

الآن تستطيع القيام بجميع العمليات الإحصائية بعد ان قمت بتغيير القيم إلى (1) و(0)، وفيما يلي نستعرض بعض هذه العمليات.

حساب الدرجة الكلية (جمع الدرجات)

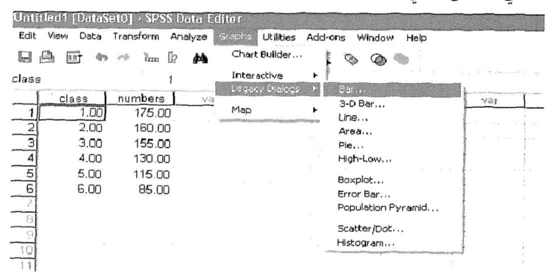
من شريط الأدوات انقر على الأمر تحوير أو Transform ومن ثم اختر الأمر Compute، في المربع على يمين صندوق الحوار تحت عنوان Target Variable أكتب اسم المتغير الجديد وليكن Total، وفي المربع أسفل Numeric Expression إختر Sum واكتب بين قوسين (q1 to q15)، ثم انقر على OK.ستجد عمود جديد في واجهة البيانات اسمه Total وهو يمثل مجموع درجات كل طالب في الامتحان.



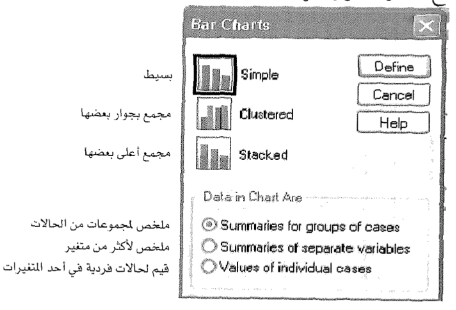
العرض البياني للدرجات

1- الأعمدة البيانية البسيطة والمزدوجة

يوفر برنامج SPSS إمكانية عرض الدرجات بطرق عديدة تتسم بالمرونة والجاذبية، ومن خلال البيانات المعروضة في المثال (1) في فصل الإحصاء، يمكن عرض اعداد الطلبة وفقا للصف الدراسي، من خلال إدخال البيانات في عمودين، بعد ذلك من Graphs نختار Legacy Dialog، ومنه نختار Bar، كما في الشكل الآتي:



سيفتح صندوق حوار عنوانه Bar Charts، نختار منه وفقا للخيارات الآتية:



نقر على Define Simple Bar، نضع عدد الطلاب في Bar Represent، والصف في Variable، بعدها ننقر على OK.

Define Simple Bar: Values of Individual Cases

Bars Represent: numbers

Category Labels

Case number

Variable: class

Panel by

Rows:

Columns:

Template

Use chart specifications from:

OK

Paste

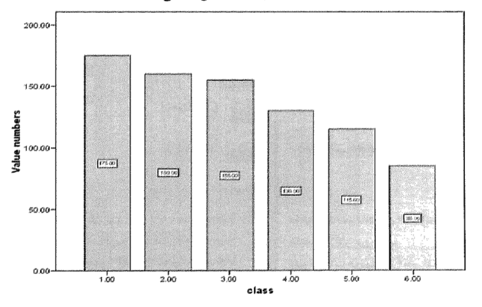
Reset

Cancel

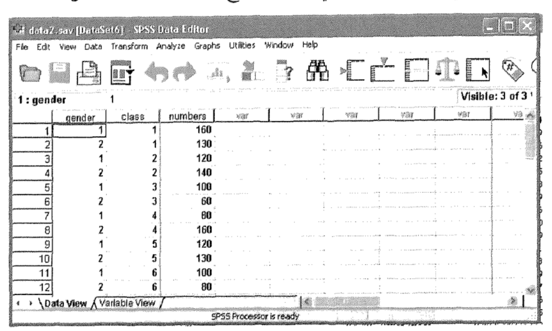
Help

Titles...

سيظهر في واجهة النتائج رسم بياني يمثل أعمدة بيانية، كل عمود يتضمن عدد الطلبة في كل صف. يمكن تحرير الرسم البياني بالنقر المزدوج عليه، ثم تغيير لون الإعمدة وعرضها، والكثير من الخصائص التي يمكن إدخالها على الرسم البياني. والشكل المبين في أدناه تم إدخال عدد الطلاب في كل عمود. يمكنك تغيير لون الأعمدة بإعطاء لون خاص بكل عمود.



ويمكن كذلك عرض الدرجات في أعمدة مزدوجة، بعدإدخال بيانات الشكل (3) في فصل الإحصاء في ملف ببرنامج SPSS كما يبين الشكل أدناه:

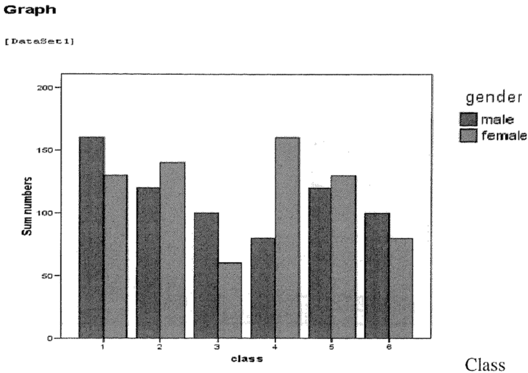


	gender	class	numbers
1	1	1	160
2	2	1	130
3	1	2	120
4	2	2	140
5	1	3	100
6	2	3	60
7	1	4	80
8	2	4	160
9	1	5	120
10	2	5	130
11	1	6	100
12	2	6	80

نفتح صندوق الحوار Bar Charts، نختار منه Clustered، سيفتح صندوق الحوار Define Clustered Bar، نقوم بإدخال عدد الطلاب فيVariable، والصف في Category Axis، ونوع الطالب في Define Clusters by، والشكل الآتي يبين ذلك:

Define Clustered Bar: Summaries for Groups of Cases
Bars Represent
N of cases
% of cases
Cum. N
Cum. %
Other statistic (e.g., mean)
Variable:
SUM(numbers)
Category Axis:
class
Define Clusters by:
gender
Panel by
Rows:
Columns:
Template
Use chart specifications from:
OK
Paste
Reset
Cancel
Help
Titles...
Options...

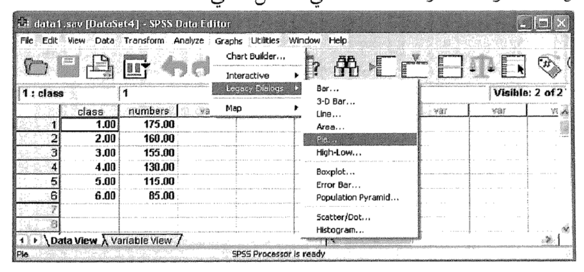
عند النقر على Ok سيظهر في واجهة النتائج الشكل البياني التالي، حاول المقارنة بين الشكل (3) فى فصل الإحصاء، والشكل أدناه.



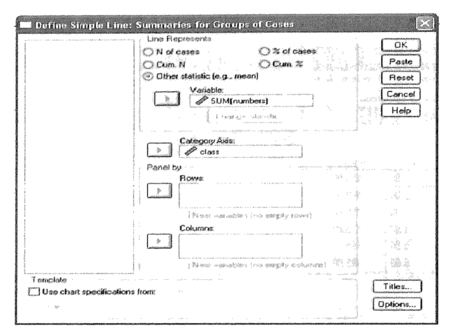
Class

(2) الخط المنكسر

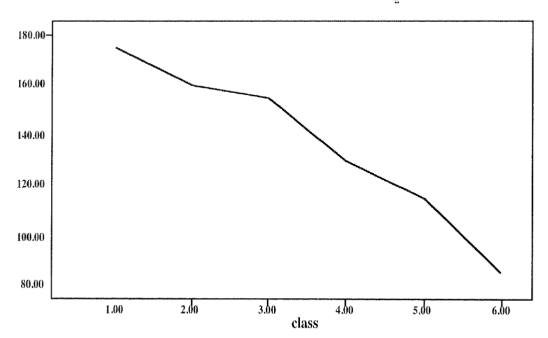
اعتمادا على بيانات المثال (1) في فصل الإحصاء، يمكن عرض أعداد الطلبة وفقا للصف الدراسي في خط منكسر. من Graphs، نختار Legacy Dialog، ومنه نختار Line، كما في الشكل الآتي:



نتقر على Line، سيفتح صندوق حوار عنوانه DefineSimple Line، نضع عدد الطلاب في Variable، والصف في Category Axis، بعدها نتقر على OK.

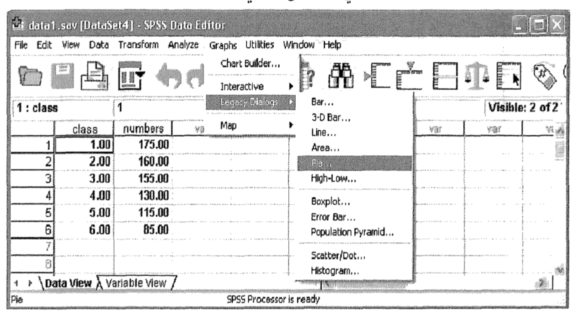


سيظهر في واجهة النتائج شكل بياني يمثل الخط المنكسر لأعداد الطلبة وفقا للصف الدراسي.



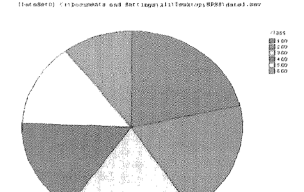
(3) الدائرة البيانية

اعتمادا على بيانات المثال (1) في فصل الإحصاء, يمكن عرض اعداد الطلبة وفقا للصف الدراسي في دائرة بيانية . من Graphs، نختار Legacy Di-alog، ومنه نختار Pie، كما في الشكل الآتي:



ننقر على Pie، سيفتح صندوق حوار عنوانه Define Simple Line، نضع عدد الطلاب في Variable، والصف في Define Slices by، بعدها ننقر على OK.

سيظهر في واجهة النتائج شكل بياني يمثل الدائرة البيانية لأعداد الطلبة وفقا للصف الدراسي.



التوزيع التكراري

ندخل بيانات المثال (2) المذكور في فصل الإحصاء، في ملف ببرنامج SPSS كما يبين الشكل ادناه :

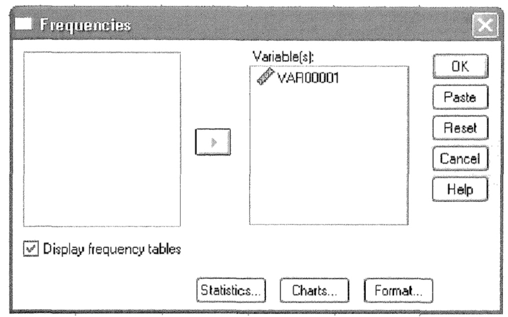
*Untitled3 [DataSet2] - SPS

File Edit View Data Transform

26 : VAR00001

	VAR00001	var
1	8.00	
2	6.00	
3	5.00	
4	9.00	
5	10.00	
6	2.00	
7	1.00	
8	.00	
9	3.00	
10	4.00	
11	7.00	
12	8.00	
13	8.00	
14	9.00	
15	8.00	
16	6.00	
17	4.00	
18	7.00	
19	3.00	
20	6.00	
21	8.00	
22	9.00	
23	7.00	
24	6.00	
25	5.00	

من Analyze نختار Descriptive Statistics، ومنها نختار Frequencies، سيظهر صندوق حوار ننقل المتغير المطلوب إلى Variable (s)، وننقر على OK



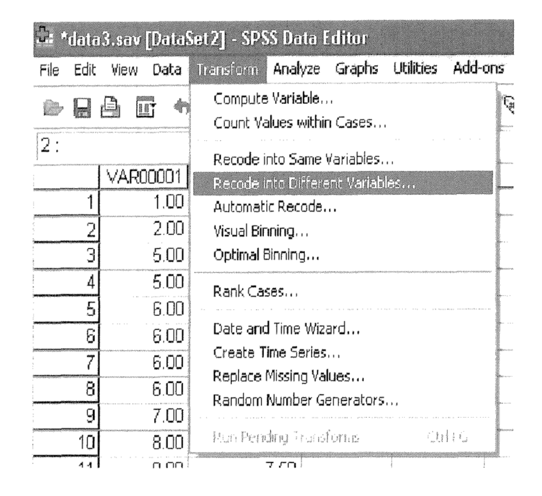
سيظهر في واجهة النتائج جدول يتضمن الدرجات الخام وتكراراتها ونسبها المئوية، والنسبة المئوية المتجمعة.

VAR00001

	Frequency	Percent	Valid Percent	Cumulative
Vaild .00	1	4.0	4.0	4.0
1.00	1	4.0	4.0	8.0
2.00	1	4.0	4.0	12.0
3.00	2	8.0	8.0	20.0
4.00	2	8.0	8.0	28.0
5.00	2	8.0	8.0	36.0
6.00	4	16.0	16.0	52.0
7.00	3	12.0	12.0	46.0
8.00	5	20.0	20.0	84.0
9.00	3	12.0	12.0	95.0
10.00	1	4.0	4.0	100.0
Total	25	100.0	100.0	

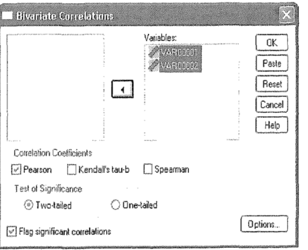
التوزيع التكراري ذو فئات

بالعودة إلى بيانات المثال (3) المذكور في فصل الإحصاء، ندخل البيانات في برنامج SPSS في عمود واحد في (50حالة)، بعد ذلك ننقر على Trans-form، ومنها نختار .Recode into Different Bariables



في Output Variable نكتب اسم المتغير الجديد وليكن newvar ثم ننقر على Change، بعدها ننقر على Old and New Values، سيظهر صندوق حوار جديد عنوانه Recode into Different Variables: Old and New Values نختار Range، ونكتب الحد الأدنى والحد الأعلى للفئة في Old Value، ثم نكتب مركز الفئة في New Value، وننقر على Add، ثم نكرر هذه الخطوة مع جميع الفئات. بعد الانتهاء من ذلك ننقر على Continue، ثم ننقر على OK.

سيظهر في واجهة البيانات عمودا جديدا اسمه newvar، وهو يمثل مركز فئة كل درجة من الدرجات الخمسين، نقوم بحساب التكرارات من Analyze نختار Descriptive Statistics، ومنها نختار Frequencies، ننقل newvar إلى Variable(s)، وننقر على OK.



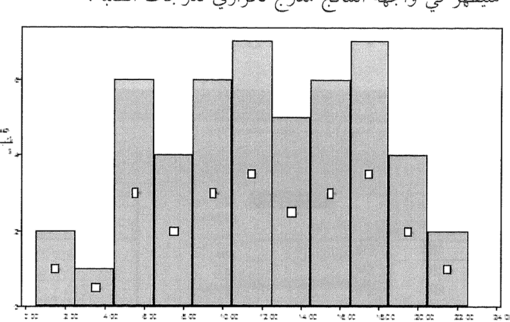
سيظهر في واجهة النتائج جدول يتضمن الدرجات الخام وتكراراتها ونسبها المئوية والنسبة المئوية المتجمعة للمتغير الجديد (مركز الفئة). يمكن مقارنة النتائج الحالية مع نتائج الجدول (3-9) في فصل الإحصاء.

1 - المدرج التكراري

من خلال بيانات المثال(3) نذهب إلى Graphs، ومنها نختار Legacy Di-alog، ثم نختار Histogram

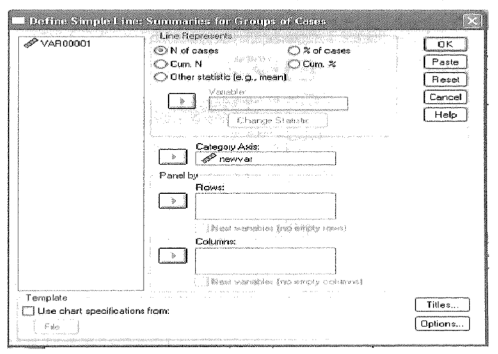
سيظهر صندوق حوار عنوانه Histogram، ننقل المتغير الجديد newvar في خانة Variable، ثم ننقر على OK.

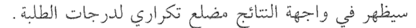
سيظهر في واجهة النتائج مدرج تكراري لدرجات الطلبة.



2 - المضلع التكراري

من خلال بيانات المثال (3) نذهب الى Graphs، ومنها نختار Legacy Dialog، ثم نختار Line سيظهر صندوق حوار نقوم بنقل المتغير الجديد newvar في Category Axis ثم ننقر على OK.





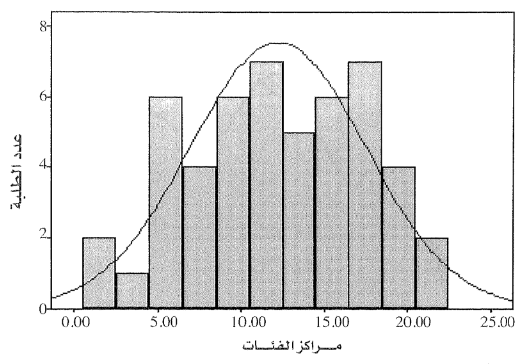
سيظهر في واجهة النتائج مضلع تكراري لدرجات الطلبة.



3 - المنحنى التكراري

إن المنحنى التكراري المتوافر في برنامج SPSS هو المنحنى الطبيعي Normal Curve؛ حيث يتيح البرنامج إمكانية المزاوجة بين توزيع درجات الطلبة والمنحنى الطبيعي، من أجل تقييم مدى قرب درجات الطلبة أو بعدها عن المنحنى الطبيعي، وطبقا لنفس الخطوات المتعلقة بالمدرج التكراري نقوم فقط بوضع إشارة (√) أمام Display Normal Curve، وننقر OK.

سيظهر في واجهة النتائج مدرج تكراري ومعه المنحنى الطبيعي لدرجات الطلبة.



مقاييس النزعة المركزية

يمكن حساب مقاييس النزعة المركزية مرة واحدة وهي المتوسط الحسابي والوسيط والمنوال، وذلك وفقا للخطوات الآتية:

من Analyze نختار Descriptive Statistics، ومنها نختار Frequencies، سيظهر صندوق حوار ننقل المتغير المطلوب حساب إحصائياته إلى خانة Variable(s)، ثم ننقر على Statistics، سيظهر صندوق حوار جديد نضع إشارة (√) أمام Mean، Median، Mode، ثم ننقر على Continue، سيتم غلق الصندوق، بعدها ننقر على OK وكما يلي (تم التطبيق على بيانات المثال الوارد في بداية هذا الفصل (ن= 20):

Frequencies
q9
q10
q11
q12
q13
q14
q15
Variable(s):
Total
OK
Paste
Reset
Cancel
Help
Display frequency tables
Statistics...
Charts...
Format...

7	14	1	0	0	1	1	0	0	1	1
8	15	1	1	1	0	1	0	1	0	1

Frequencies: Statistics
Percentile Values
Quartiles
Cut points for: 10 equal groups
Percentile(s):
Add
Change
Remove
Central Tendency
Mean
Median
Mode
Sum
Continue
Cancel
Help
Values are group midpoints
Dispersion
Std. deviation
Minimum
Variance
Maximum
Range
S.E. mean
Distribution
Skewness
Kurtosis

سيظهر في واجهة النتائج جدول فيه مؤشرات النزعة المركزية الثلاثة.

Statistics

Total

N	Valid	20
	Missing	0
Mean		8.60
Median		9.50
Mode		10

ملاحظة على مقاييس النزعة المركزية

1- يمكن تطبيق الخطوات نفسها على بيانات الامثلة الواردة في فصل الإحصاء، والمقارنة بين نتائج برنامج SPSS، ونتائج الحساب اليدوي.

2- يمكن حساب مقاييس النزعة المركزية في برنامج SPSS باكثر من طريقة، فمثلا المتوسط الحسابي نجده في معظم التحليلات الإحصائية.

مقاييس التشتت

يمكن حساب مقاييس التشتت بنفس الخطوات التي تمت في حساب مقاييس النزعة المركزية، ولكن في هذه المرة تضع الإشارة أمام Std deviation، و Variance، و Range.

سيظهر في واجهة النتائج جدول يتضمن مقاييس التشتت :

Statistics

Total

N	Valid	20
	Missing	0
Std. Deviation		3.136
Variance		9.832
Range		12
Minimum		2
Maximum		14

ملاحظة على مقاييس النزعة المركزية ومقاييس التشتت
يمكن حساب جميع مقاييس النزعة المركزية وجميع مقاييس التشتت مرة واحدة وفقا للخطوات التي تمت الإشارة إليها سابقا، وبذلك فإذا كان المعلم يخطط لحسابها جميعاً يمكنه ذلك بدلا من تكرار الخطوات، وفيما يأتي مثالاً لذلك:

Statistics

Total

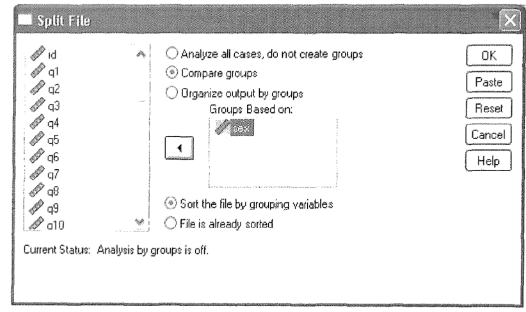
N	Valid	20
	Missing	0
Mean		8.60
Median		9.50
Mode		10
Std. Deviation		3.136
Variance		9.832
Range		12
Minimum		2
Maximum		14

معرفة متوسط كل مجموعة على حدة (تجزئة الملفات)

بما أنه لدينا مجموعة من الطلبة تضم ذكورا وإناثا، فمن المناسب معرفة متوسط كل مجموعة على حدة، ولعمل ذلك اختر Split file من قائمة Data. في صندوق الحوار الذي يظهر أمامك اختر Compare Groups وادخل اسم

المتغير الذي تود تقسيم الطلبة به في المربع، في هذه الحالة نقسم الملف حسب متغير النوع (Sex) .

ملاحظة :
كل التحليلات الإحصائية التي ستجري على الملف ستكون وفق كل جنس على حدة ما لم نوقف (Split file)



من أجل السرعة في تنفيذ المطلوب، ننقر على الايقونة الرابعة Dialog Recall، ونختار Descriptives، سيظهر صندوق حوار ننقر فيه على OK) لأن المتغير المطلوب حساب مؤشراته الإحصائية موجود أصلا في المكان المطلوب).

1 : Id

Split File
Descriptives
Reliability Analysis
Compute Variable
Bivariate Correlations
One-Way ANOVA
Independent-Samples T Test
Frequencies
Recode Into Same Variables
Chart Builder
Define Multiple Line: Summaries of Separate Variables
Create Bar Chart

	id	sex	q10	q11
1	1	1	1	1
2	2	1	1	0
3	3	1	1	1
4	7	1	1	1
5	9	1	1	1
6	10	1	1	0
7	14	1	1	0

تظهر النتائج على شكل جدول فيه المؤشرات الإحصائية لنتائج الذكور وهي منفصلة عن نتائج الإناث، فمثلا نجد أن أقل درجة كانت لدى الذكور (2)، وان أعلى درجة كانت لدى الإناث (14)، وهكذا نستطيع أن نقارن في جوانب أخرى كالمتوسط والانحراف المعياري.

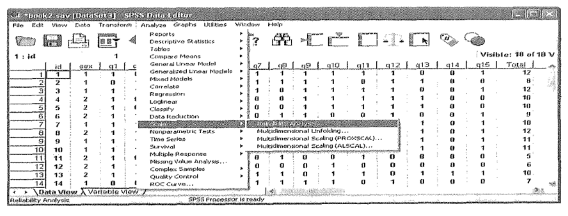
Descriptives

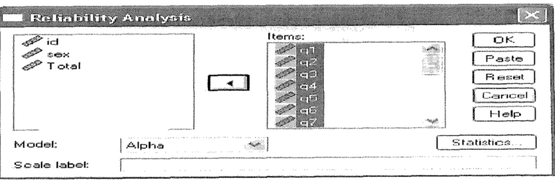
Descriptive Statistics

sex		N	Minimum	Maximum	Mean	Std. Deviation
male	Total	10	2	12	8.60	3.204
	Valid N (listwise)	10				
female	Total	10	4	14	8.60	3.239
	Valid N (listwise)	10				

حساب ثبات الاختبار

يمكن حساب معامل ثبات الاختبار، بالنقر على Analyze، ثم ننقر على Scale، وبعدها نختار Reliability سيظهر صندوق حوار عنوانه Reliability Analysis ننقل جميع اسئلة الاختبار (15 سؤالاً) إلى المربع المعنون Items، ومن ثم ننقر على OK.





ستظهر في واجهة النتائج معلومات تتعلق بعدد فقرات الاختبار، ومعامل ألفا- كرونباخ، وفي هذا المثال قيمة معامل الثبات 0.708 وهي قيمة مرتفعة في اختبارات التحصيل.

Reliability Statistics

Cronbach's Alpha	N of Items
.708	15

ويمكن تحسين معامل الثبات من خلال حساب العلاقة بين الدرجة على كل سؤال والدرجة الكلية على الاختبار، واستبعاد الأسئلة التي ترتبط بعلاقة سالبة مع المجموع الكلي، وحذف الأسئلة التي ترتبط ارتباطا ضعيفا بالدرجة الكلية.

Reliability Analysis: Statistics

Descriptives for	Inter-Item	
☐ Item	☐ Correlations	Continue
☐ Scale	☐ Covariances	Cancel
☑ Scale if item deleted		Help
Summaries	**ANOVA Table**	
☐ Means	◉ None	
☐ Variances	○ F test	
☐ Covariances	○ Friedman chi-square	
☐ Correlations	○ Cochran chi-square	
☐ Hotelling's T-square	☐ Tukey's test of additivity	
☐ Intraclass correlation coefficient		
Model: Two-Way Mixed	Type: Consistency	
Confidence interval: 95 %	Test value: 0	

يتم حساب العلاقة بين الدرجة على كل سؤال والدرجة الكلية على الاختبار، من خلال النقر على Statistics الموجودة في صندوق Reliability Analysis، بعدها نضع إشارة (√) أمام Scale if item deleted، وننقر على Continue، وبعدها OK.

سيظهر في واجهة النتائج جدول يتضمن أربعة معلومات (الأرقام الموضحة في حال حذف السؤال) هي: المتوسط الحسابي، والتباين، ومعامل الارتباط، ومعامل ألفا- كرونباخ. ويلاحظ ارتباط السؤال الرابع ضعيف (0.166)، وارتباط السؤال الثامن سالب (-0.100)، وارتباط السؤال الرابع عشر ضعيف (0.034).

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
q1	7.75	9.250	.201	.705
q2	8.05	8.471	.370	.687
q3	8.05	8.787	.259	.701
q4	7.95	9.103	**.166**	.711
q5	8.05	8.471	.370	.687
q6	7.90	8.832	.279	.698
q7	8.05	8.155	.486	.671
q8	8.00	9.895	**-.100**	.742
q9	8.00	9.053	.174	.711
q10	7.75	8.934	.348	.692
q11	8.15	8.345	.416	.681
q12	7.90	8.095	.567	.663
q13	8.15	8.134	.494	.670
q14	8.55	9.734	**.034**	.714
q15	8.10	7.884	.585	.658

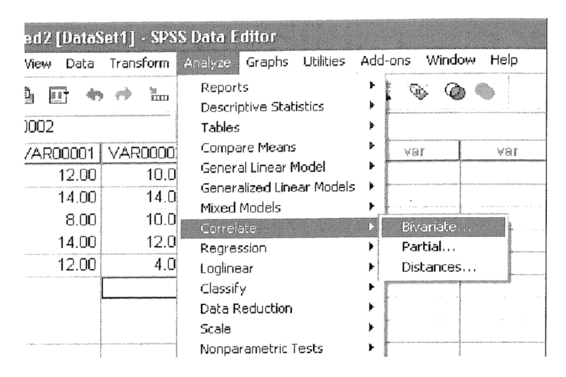
باستبعاد هذه الاسئلة من التحليل سوف يرتفع معامل الثبات. يتم ذلك بإعادة خطوة حساب معامل الثبات، واستبعاد الأسئلة الثلاث (4، 8، 14). سيبقى لدينا 12 سؤالاً فقط. ومعامل ثباتها 0.755، وهو أعلى من معامل الثبات السابق الذي كان 0.708 وعدد الأسئلة 15.

Reliability Statistics

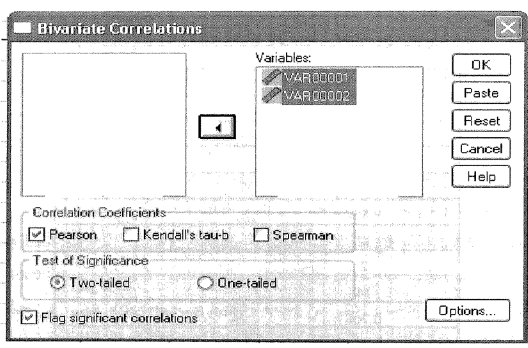
Cronbach's Alpha	N of Items
.755	12

مقاييس العلاقة

لحساب العلاقة بين متغيرين، نذهب إلى Analyze، ونختار Correlate، ومنها نختار Bivariate (سنطبق بيانات المثال 14 في فصل الإحصاء).



في صندوق الحوار Bivariate Correlations نضع المتغيرات المطلوب حساب معامل الارتباط لها في خانة Variables، والبرنامج معد لحساب معامل ارتباط بيرسون، ولفرضية بذيلين، ويستطيع المستخدم تغيير معامل الارتباط باختيار معامل كندال- تاو، أو معامل سبيرمان للرتب.



ستظهر في واجهة النتائج قيمة معامل الارتباط، وهي 0.327 وعلى الرغم من أن العلاقة أيجابية بين درجات الطلبة في الرياضيات ودرجاتهم في الفيزياء، إلا أنها غير دالة إحصائيا لان مستوى الثقة 0.591 (وهو يشير إلى نسبة خطأ غير مقبولة، ويفترض أن تكون نسبة الخطأ 0.05 أو أقل من ذلك).

Correlations

[DataSet1]

Correlations

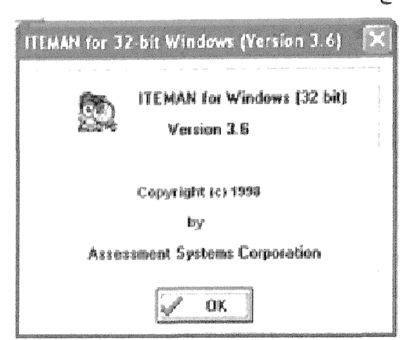
		VAR00001	VAR00002
VAR00001	Pearson Correlation	1	.327
	Sig. (2-tailed)		.591
	N	5	5
VAR00002	Pearson Correlation	.327	1
	Sig. (2-tailed)	.591	
	N	5	5

ملخص عام

بعد الانتهاء من حساب التحليلات الإحصائية لدرحات الطلبة في اختبار ما (وليكن درجات المثال المعروض في بداية هذا الفصل) نلخص أهم تلك النتائج في الجدول الآتي:

عدد الأسئلة	15
عدد الطلبة	20
المتوسط الحسابي	8.60
الوسيط	9.50
المنوال	10
الانحراف المعياري	3.36
التباين	9.832
المدى	12
أقل درجة	2
أعلى درجة	14
معامل الثبات (ألفا - كرونباخ)	0.708
معامل الثبات الجديد (12 سؤالاً)	0.755

ثانياً - برنامج ITEMAN

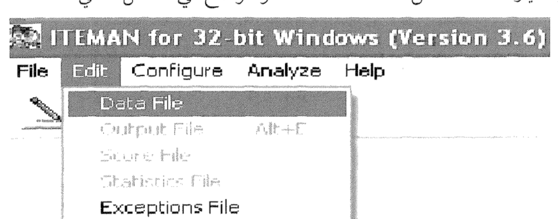


مقدمة

يُعد برنامج (ITEMAN (Item Analysis من البرامج السهلة الاستخدام لغرض تحليل فقرات الاختبارات التحصيلية بشكل خاص والإستبانات البحثية بشكل عام. فهو يعطي مؤشرات إحصائية سواء على مستوى الفقرة الواحدة أو على مستوى الاختبار ككل. فعلى مستوى الفقرة يقوم البرنامج بحساب معامل صعوبة الفقرة ومعامل تمييزها (مدى إسهام الفقرة في معامل ثبات الاختبار). وعلى مستوى الاختبار ككل، يعطي البرنامج المؤشرات الإحصائية التالية: المتوسط والانحراف المعياري، ومعامل الثبات، والتفلطح، والانحدار، وغيرها. كما توجد بالبرنامج خاصية استخراج ملف إحصاءات منفصل للاستخدام في تحليلات إحصائية متقدمة. وما يميز البرنامج ايضا هو إمكانية فتح ملف مدخلات input file يسمى بملف الاستثناءات exception file يمكنك من إدخال أكثر من إجابة واحدة صحيحة لأي فقرة في حالة اكتشافك وجود ذلك في الفقرة.

خطوات عمل البرنامج:

لكي تكوّن ملف بيانات جديد أو تفتح أحد الملفات القديمة، عليك باختيار Data file من القائمة Edit كما هو موضح في الشكل الآتي:



مكونات ملف البيانات:

يتكون ملف البيانات في برنامج ITEMAN، من خمسة عناصر:

1- سطر التحكم الذي يعطي وصفا للبيانات.

2- سطر مفتاح الإجابة.

3- سطر يبين عدد الخيارات لكل فقرة من فقرات الاختبار.

4- سطر يوضح الفقرات التي سوف يشملها التحليل.

5- بيانات الطلاب (إجابات الطلاب على الامتحان)، بحيث يكون لكل حالة (مستجيب، أو مفحوص) سطر واحد. وفيما يلي وصفا لكل عنصر على حدة:

1 - سطر التحكم:

السطر الاول من ملف البيانات يتضمن معلومات عن الاختبار، وهو من أهم المعلومات في برنامج ITEMAN، وأي خطأ فيه يؤثر على النتائج والمعلومات التي يتضمنها هي: عدد فقرات الإختبار، والرمز الذي يشير إلى الإجابة المفقودة، والرمز الذي يشير إلى عدم تمكن الطالب من الوصول إلى السؤال بسبب عامل الوقت، وأخيرا هوية المفحوص. وفيما يأتي تفاصيل ذلك.

العمود	البيـــــانات
3-1	عدد فقرات الإختبار(بحد أقصى 750 فقرة لكل مفحوص)
4	فاضي (مسافة)
5	رمز الإجابات المفقودة omitted (رقم أو حرف)
6	فاضي (مسافة)
7	رمز الأسئلة التي لم يصلها المفحوص (not reached رقم أو حرف)
8	فاضي (مسافة)
10-9	هوية المفحوص ID، بحد أقصى 80 رمزا.

2 - مفتاح الإجابة:

السطر الثاني في ملف البيانات يأخذ شكلين مختلفين: الأول عندما تكون البيانات ثنائية التصحيح Dichotomous (صفر، او واحد) كما هو الحال في اختبارات التحصيل؛ أما الشكل الثاني فعندما تكون البيانات متعددة الأجوبة (ليس بالضرورة هناك إجابة صحيحة واحدة) كما هو الحال في سلالم التقدير Likert scale مثلا.

ففي حالة التعامل مع بيانات إختبار تحصيلي، أو اختبار قدرات فهذا السطر يمثل مفتاح إجابة الأختبار. بحيث أن الرمز الأول يمثل الإجابة الصحيحة للسؤال الأول، والرمز الثاني يمثل الإجابة الصحيحة للسؤال الثاني إلخ .....؛ هذا ويمكن مع ITEMAN استخدام الأرقام من 1-9 أو الحروف من A-I .

أما في حالة التعامل مع بيانات استبيان، فهذا السطر يبين الفقرات الموجبة والفقرات السالبة. فيرمز للفقرات الموجبة بالرمز (+)، وفي هذه الحالة يعطي الرقم (1) أو الحرف (A) قيمة محددة هي (1) والرقم (2) أو الرمز (B) القيمة (2) وهكذا. ويرمز للفقرات السالبة بالرمز (-) وفي هذه الحالة تعكس القيم بالنسبة لهذه الفقرات.

كما يمكن التعامل مع بيانات متعددة الإجابة (ثنائية، ثلاثية، ومتعددة)، كأن يكون جزء من البيانات من نوع اختيار من متعدد، والجزء الثاني هو أسئلة مفتوحة، والجزء الأخير أسئلة لها مقياس متدرج (مثلاً: Liker scale)

3 - عدد الخيارات:

السطر الثالث يبين عدد الخيارات (أو البدائل) لكل سؤال (وهذه مهمة جدا، عندما يكون عدد الخيارات في فقرات الإختبار غير ثابت، مثلا: سؤال اختيار من متعدد من أربعة بدائل، وسؤال آخر اختيار من متعدد من ثلاثة بدائل).

4 - رموز الفقرات المتضمنة في التحليل:

السطر الرابع يبين الفقرات التي ترغب في تضمينها، وتلك التي لا ترغب في تضمينها في التحليل الإحصائي، بحيث تأخذ الفقرات التي ترغب في تضمينها الرمز Y وتعني (Yes) وتأخذ الفقرات التي لا ترغب في تضمينها في التحليل الرمز N وتعني .(No)

5 - إدخال البيانات

تبدأ البيانات من السطر الخامس؛ بحيث يتم إدخال بيانات كل حالة (أي كل طالب او مستجيب) في سطر منفصل بغض النظر عن عدد الفقرات، على ان يبدأ السطر بهوية المستجيب (عدد الرموز، ويجب ان يكون مساويا للعدد الذي تم تحديده في سطر التحكم، وهو السطر الأول من الملف).

مثال تطبيقي (1):

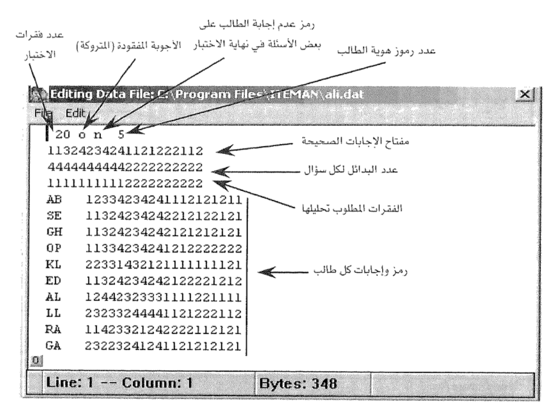
تم ادخال بيانات عشرون طالباً في برنامج ITEMAN وفقا للخطوات التي تم عرضها سابقا، وفيما يلي شرحا للعملية والخطوات التي في ضوءها ادخال البيانات:

- يبين السطر الأول عدد مفردات الاختبار 15 مفردة، والرمز المستخدم في حالة الإجابة المفقودة هو (o)، أما الرمز المستخدم في حالة عدم

الطالب من الإجابة على بعض الأسئلة بسبب عامل الوقت هو (n)، وأما رمز هوية الطالب فهو (2) أو رقمين.

- يبين السطر الثاني مفتاح إجابة الاختبار.
- يبين السطر الثالث عدد الخيارات لكل سؤال.
- يبين السطر الرابع مدى رغبتنا في تضمين كل سؤال في عمليات التحليل المختلفة.

وبدءاً من السطر الخامس تبدأ إجابات الطلاب (كل سطر يبدأ برقم الطالب من 1-20 متبوعا بـ 15 رمزا تمثل إجاباته على الـ 15 سؤالا في الامتحان).



حفظ البيانات

بعد الانتهاء من ادخال البيانات، يتعين حفظها، لأجل ذلك إختر من قائمة File الأمر Save as ثم إدخل إسم الملف، ولكن بشرط أن ينتهي أسم الملف بنقطة والكلمة (Dat)؛ لكي يستطيع البرنامج التعرف عليه. ومثال على ذلك سنختار كلمة (test.dat) لاسم الملف.

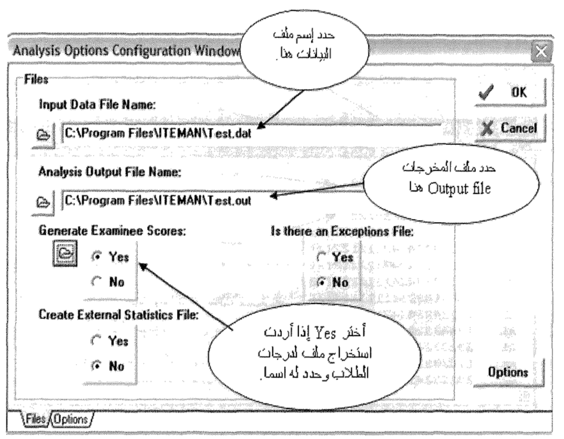
تحليل البيانات

قبل تحليل البيانات، يتعين تعريف البرنامج بطريقة ترتيب البيانات Con-figuration، وهذا الاجراء يتم من خلال الخطوات الآتية:

1 - أغلق ملف البيانات بعد ان تتأكد من حفظه.

2 - من الواجهة الرئيسية إختر قائمة Configure، وإضغط على Alt +C.

- في صندوق الحوار التالي حدد اليبانات المطلوبة، ومن ثم أنقرعلى الزر OK :



لتحليل البيانات، إضغط على Alt +A، أو انقر على Analyze سيقوم البرنامج بضغطه زر واحدة بإجراء كافة التحليلات الإحصائية للاختبار. ولمعاينة ملف المخرجات، اضغط على Alt+ E، يشتمل ملف النتائج على معلومات اولية يمكن استخدامها للتحقق من صحة إدخال البيانات. والشكل التالي يفسر بعض إجزاء ملف المخرجات بالنسبة للاختيار من متعدد.

معامل الصعوبة — مستوى التمييز

```
Browsing ... File: C:\Program Files\ITEMAN\Test.out
File  Edit

                 Item Statistics                    Alternative Statistics
          ---------------------------------   ------------------------------------
Seq.  Scale   Prop.    Disc.   Point              Prop. Endorsing      Point
No.   -Item   Correct  Index   Biser.      Alt.   Total   Low   High   Biser.  Key
----  -----   -------  -----   ------      -----  -----   ----  ----   ------  ---

  1    0-1     .85      .29     .34          1     .05    .14    .00    -.33
                                             2     .85    .71   1.00     .34    *
                                             3     .10    .14    .00    -.16
                                             4     .00    .00    .00
                                         Other     .00    .00    .00

  2    0-2     .55      .69     .54          1     .55    .14    .83     .54    *
                                             2     .20    .43    .17    -.24
                                             3     .25    .43    .00    -.39
                                             4     .00    .00    .00
                                         Other     .00    .00    .00

Line: 1 -- Column: 1        Bytes: 11435
```

كما يعطينا البرنامج في هذا الملف العديد من الاحصاءات، نذكر منها:

Variance	9.55	التباين
Std. Dev.	3.09	الانحراف المعيارى
Skew	-0.198	الالتواء
Kurtosis	-0.744	التفلطح
Minimum	2	أدنى درجة
maximum	14	اعلى درجة
Midian	9	الوسيط
Alpha	0.711	ألفا - كرونباخ

وبما أنه كان من ضمن خياراتنا عند التحليل استخراج ملف لنتائج الطلاب، فلمعاينة هذا الملف إختر Score File من قائمة Edit. والشكل التالي يوضح ملف نتائج الطلاب.

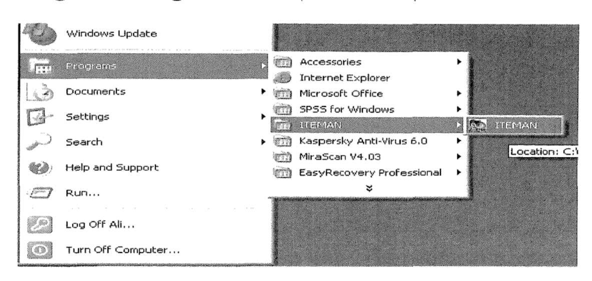
```
Browsing Score File: C:\Program ...
File  Edit
     2    1    Scores for examinees
  01   11.00
  02    8.00
  03   12.00
  04   10.00
  05   10.00
  06    9.00
  07   10.00
  08   13.00
  09   11.00
  10   11.00
  11    5.00
  12    6.00
  13    9.00
  14    6.00
  15    7.00
  16    6.00
  17    4.00
  18    6.00
  19   14.00
  20    2.00
```

والملاحظ على ملف النتائج أنه مرتب بشكل مبسط، ويمكن ان يدخل مباشرة في سجل المقرر، خاصة إذا ما أدخلنا أسماء الطلاب في ملف البيانات.

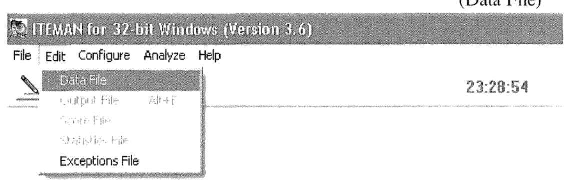
مثال تطبيقي (2):

لنفرض اننا طبقنا اختبار موضوعي على (10) طلاب، يتالف الاختبار من (20) سؤالاً. (10) أسئلة من نوع اختيار من متعدد لها أربعة بدائل، و(10) أسئلة من نوع الصواب والخطأ. ولأجل اجراء التحليل الأحصائي لأجابات الطلبة وفقا لبرنامج ITEMAN (نفتح ملف جديد) وفقا للخطوات الآتية:

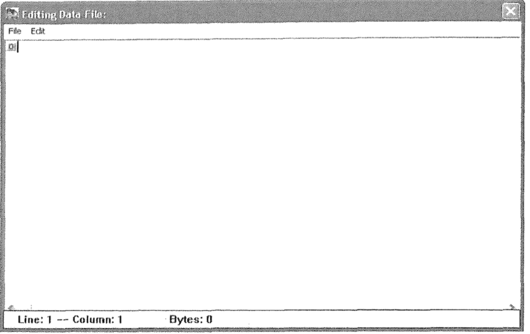
1 - ننقر على start، ثم Programs، ثم ITEMAN ستفتح واجهة البرنامج.



2 - عند ظهور واجهة البرنامج على الشاشة، ننقر على Edit، ومنها نختار (Data File)



ستظهر واجهة البيانات المسماة (Editing Data File)



3 - على شاشة البيانات المذكورة أعلاه أبدأ بإدخال المعلومات الخاصة بملف الاختبار، وهي 4 معلومات، وكما يأتي:

السطر الأول: نترك مسافة واحدة ونكتب عدد أسئلة الاختبار (20)، ثم نترك مسافة واحدة، ونحدد الأجابات المحذوفة أو المفقودة، وفي حالة عدم وجودها نكتب حرف (o)، ثم نترك مسافة، ونكتب حرف (n) للدلالة على عدم تمكن الطالب من اجابة الفقرات الأخيرة بسبب عدم كفاية الوقت المخصص للأختبار (أي انتهى وقت الاختبار ولم ينتهي الطالب من الاجابة)، ثم نترك مسافة واحدة إذا كان عدد الرموز من (10-80) رمزا أو حرفاً، أو نترك مسافتين إذا كان عدد الرموز من (1-9) رموز أو حروف، ولنحدده بالحرفين الأولين من أسمه مثلاً، مع ترك فراغ بمقدار ثلاث مسافات، وبهذا سيكون عدد رموز هوية الطالب (5) رموز. ستكون كل هذه المعلومات مكتوبة على السطر الأول (من اليسار الى اليمين)، كما يلي: 20 o n 5

السطر الثاني: ننتقل إلى السطر الثاني بالضغط على Enter، ثم نكتب في

هذا السطر وبالتسلسل (من السؤال الأول) أرقام الاختيارات الصحيحة لكل سؤال. مثال ذلك: السؤال الأول اختياره الصحيح (1)، والثاني (1)، والثالث (3)، والرابع (2)، والخامس (4) وهكذا مع باقي أسئلة الاختيار من متعدد، وأما أسئلة الصواب والخطأ فإن الاجابة الصحيحة هي (1) للسؤال الحادي عشر، و(1) للثاني عشر، و(2) للثالث عشر، وهكذا مع باقي أسئلة الصواب والخطأ. وفي مثالنا نكتب الأجابات الصحيحة لعشرين فقرة، وبدون ترك مسافات كالأتي: 11324234241121222112

السطر الثالث: نتقل إلى السطر الثالث بالضغط على Enter، ثم نكتب فيه (عدد بدائل كل سؤال، بنفس عدد وترتيب الاسئلة، اسئلة الاختيار من متعدد، ثم أسئلة الصواب والخطأ، وفي هذا المثال عدد بدايل الاختيار من متعدد (4)، وعدد بدائل الصواب والخطأ (2)، وعليه سنكتب في هذا السطر: 44444444442222222222

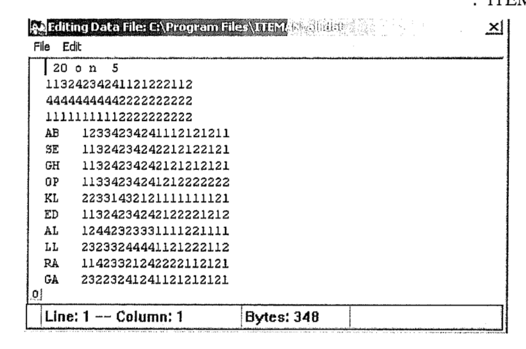
السطر الرابع: ويتضمن تحديد الاسئلة التي نرغب في تضمينها بعملية التحليل، والأسئلة التي لا نرغب تضمينها في التحليل الاحصائي، بحيث تأخذ الأولى الرمز Y، والثانية تأخذ الرمز N ويمكن أن يقوم البرنامج يتحليل أسئلة الاختيار من متعدد واسئلة الصواب والخطأ بشكل منفصل، وفي مثالنا هذا سنعطي الرقم 1 لاسئلة الاختيار من متعدد، والرقم 2 لأسئلة الصواب والخطأ، وعليه سنكتب: 11111111112222222222.

السطر الخامس: في هذا السطر نبدأ بإدخال رمز الطالب أولا (هويته أو اسمه ID)، ثم إجاباته على أسئلة الاختبار كما وردت في أوراق الإجابة مع ملاحظة أن رمز الطالب يجب أن لا يزيد عن العدد الذي تم تحديده في السطر الأول وهو (2).

ومن المثال نكتب إجابات الطلبة كما يأتي (علما بأن عدد الطلبة غير محدد أي يمكن إدخال عدد غير محدد من الطلبة)، ولكن في المثال العدد هو (10) طلاب.

AB 12334234241112121211
SE 11324234242212122121
GH 11324234242121212121
OP 11334234241212222222
KL 22331432121111111121
ED 11324234242122221212
AL 12442323331111111111
LL 23233244442222222222
RA 11423321242222112121
GA 23223241241121212121

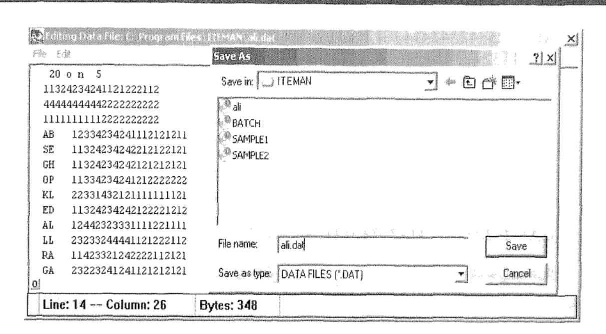
وفيما يلي توضيح لمدلول الارقام التي يتم ادخالها في واجهة بيانات : ITEMAN



حفظ البيانات:

الخطوة الأولى: بعد الانتهاء من الخطوة الخاصة بإدخال البيانات (إجابات الطلاب) نبدأ بحفظ البيانات كما يلي:

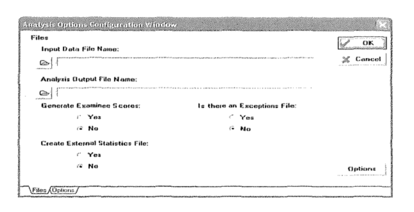
- انقر على File، ومنه انقر على Save As، فيظهر صندوق حوار اسمه Save As
- في المكان المخصص لاسم الملف، اكتب رمز أو كلمة تميز بها الملف الحالي مثل اسمك dat+. ويمكن الأكتفاء بكتابة الأسم.
- ثم اضغط على Save



الخطوة الثانية: أغلق واجهة البيانات لتعود إلى الصفحة الرئيسية للبرنامج وهي: (Iteman for 32 - bit Windows)

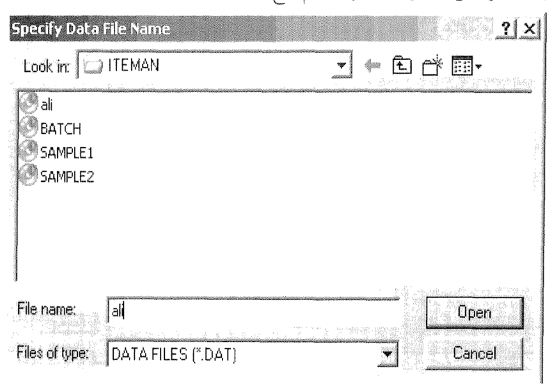
ومن هذه الصفحة اضغط على كلمة Configure فتظهر تحتها كلمة: (Go Alt+c)، اضغط عليها، أو انقر على صورة القلم بدلا عن ذلك، فيظهر لك صندوق جديد اسمه (Analysis Options).

20 48 25



الخطوة الثالثة: عندما يظهر الصندوق الجديد وهو Analysis Options انقر ايقونة صفراء عنوانها Input Data لاختيار الملف المطلوب وتحديد مكان تخزينه.

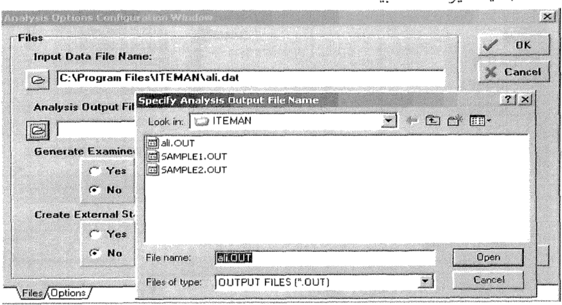
بعد النقر على الايقونة الصفراء سيتم فتح نافذة .Specify Data File Name



الخطوة الرابعة: من صندوق الحوار السابق اكتب اسم الملف المطلوب، أو حدد الملف، وانقر على (Open) سيتم غلق الصندوق، وسيظهر الصندوق السابق Analysis Options .

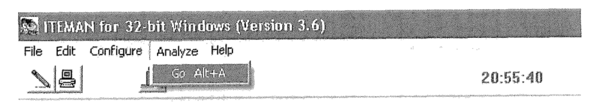
Analysis Options Configuration Window
Files
Input Data File Name:
C:\Program Files\ITEMAN\ali.dat
OK
Cancel
Analysis Output File Name:
Generate Examinee Scores:
Yes
No
Is there an Exceptions File:
Yes
No
Create External Statistics File:
Yes
No
Options
Files Options

الخطوة الخامسة: اضغط على Analysis Output سيظهر صندوقا جديدا، حدد فيه اسم الملف الذي ترغب ان تضع نتائج التحليل فيه. وعادة ما يكون ملف جديد غير ملف البيانات.



الخطوة السادسة: في صندوق الحوار Analysis Options اختر Yes , No حسب حاجتك للاحصاءات الوصفية، ثم انقر OK ستعود إلى الواجهة الرئيسة للبرنامج وهي: Iteman for 32 - bit

الخطوة السابعة (الأخيرة): من الواجهة الرئيسة للبرنامج انقر على Ana- lyze، ستظهر تحتها عبارة (Go Alt+A)،



اضغط عليها، سيظهر لديك نفس الصندوق وفيه مجموعة عبارات، الاولى تخص تحليل البيانات، والثانية لاعطاء الأمر بطباعة النتائج، والثالثة للخروج من البرنامج. اضغط على الايقونة View out put، الموجودة تحت كلمة Configure، أو أنقر على Edit ومنها على Out put File ستظهر صفحة التحليلات الإحصائية النهائية، والمتضمنة جميع البيانات الخاصة بالتحليل الإحصائي للاختبار، ويمكن ملاحظة المؤشرات الاحصائية لجميع فقرات الاختبار (من الفقرة 1 إلى الفقرة 20) وهو عدد الفقرات في مثالنا.

ومن المعلومات العديدة عن الأختبار والفقرات انظر إلى الآتي:

العمود الثالث من اليسار: مستوى صعوبة الفقرة (Prop. Correct)

العمود الرابع من اليسار: مستوى تمييز الفقرة (Disc. index)

العمود الخامس من اليسار: معامل ارتباط الفقرة بالدرجة الكلية للأختبارPoint Biserial Corr.

وهنالك معلومات اخرى مفيدة في أسفل الصفحة من جهة اليسار منها:

- عدد فقرات الاختبار N. of items
- عدد الطلاب الممتحنين N. of Examinees
- الوسط الحسابي للعلامات . Mean
- التباين. Variance
- الانحراف المعياري ..St. Dev
- الوسيط . Median
- معامل الفا .Alpha

وهناك رسم بياني يتضمن النسب المئوية لأجابات الطلاب الصحيحة على الفقرات ومعلومات اخرى.

طباعة نتائج التحليل:

للحصول على نسخة مطبوعة من المعلومات الخاصة بأحصائيات الفقرة والأختبار (الجداول الأحصائية)، أنقر على كلمة File ومنه انقر على+ Print

Alt + P فتحصل على نسخة مطبوعة من الأحصائيات. وفيما يلي نتائج ITEMAN للمثال الثاني.

الصفحة الاولى:

```
                  ITEMAN (tm ) for32-bit Windows  . Version3.6        Page 1
        Copyright (c )1982 - 1998              by Assessment Systems

                   Conventional Item and Testlysis Program


 Item analysis for data from file :Program Files\ITEMAN\ali.dat
Date :May 16 .2008                                         Time :0:48 AM


    ********************                 ANALYSIS SUMMARY     **********


           Data (Input )File  :C:\Program Files\ITEMAN\ali.dat
            Analysis Output   :C:\Program Files\ITEMAN\ali.OUT
          Score Output File   :C:\Program Files\ITEMAN\students.SCR
             Exceptions File  :NONE
          Statistics Output   :C:\Program Files\ITEMAN\statistics.PAR

        Scale Definition  :  DICHOT =Dichotomous   MPOINT =Multipoint /Survey


  Scale :            1        2
                  ------- -------
      Type of Scale   DICHOT
 N of Items          10       10
 N of Examinees      10       10


                     *****            CONFIGURATION   *****

                           Type of  : Point-Biserial

                    Correction for  : NO

                  Ability Grouping  : YES

                 Subgroup Analysis  : NO

         Express Endorsements As  : PROPORTIONS

            Score Group Interval  :  1
```

ملف البيانات

ملف نتائج

ملف نتائج الطلبة

ملف الإحصاءات

الصفحة الثانية:

ITEMAN (tm) for 32-bit Windows Version 3.6 Page 2
Copyright (c) 1982 - 1998 by Assessment Systems

Conventional Item and Test Analysis

Item analysis for data from file : \Program Files\ITEMAN\als.dat
Date May 16 2006 Time : 0:46 AM

		Item Statistics			Alternative					
Seq. No.	Scale -	Prop. Item	Disc. Correct	Point Index	Alt.	Prop	Endorsing Total	Low	Point High	Key
1	1-1	.70	.50	.62	1	.70	.50	1.00	.62	*
					2	.30	.50	.00	-.62	
					3	.00	.00	.00		
					4	.00	.00	.00		
					Other	.00	.00	.00		
2	1-2	.50	1.00	.76	1	.50	.00	1.00	.76	*
					2	.30	1.00	.00	-.49	
					3	.20	.00	.00	-.40	
					4	.00	.00	.00		
					Other	.00	.00	.00		
3	1-3	.60	.50	.71	1	.00	.00	.00		
					2	.20	.00	.00	-.40	
					3	.60	.50	1.00	.71	*
					4	.20	.50	.00	-.47	
					Other	.00	.00	.00		
4	1-4	.50	1.00	.47	1	.00	.00	.00		
					2	.50	.00	1.00	.47	*
					3	.40	.50	.00	-.17	
					4	.10	.50	.00	-.51	
					Other	.00	.00	.00		
5	1-5	.50	1.00	.94	1	.10	.50	.00	-.41	
					2	.10	.50	.00	-.51	
					3	.30	.00	.00	-.42	
					4	.50	.00	1.00	.94	*
					Other	.00	.00	.00		
6	1-6	.70	1.00	.68	1	.00	.00	.00		
					2	.70	.00	1.00	.68	*
					3	.20	.50	.00	-.47	
					4	.10	.50	.00	-.41	
					Other	.00	.00	.00		
7	1-7	.60	.50	.71	1	.00	.00	.00		
					2	.20	.50	.00	-.47	
					3	.60	.50	1.00	.71	*
					4	.20	.00	.00	-.40	
					Other	.00	.00	.00		

الصفحة الثالثة:

ITEMAN (tm) for32-bit Windows Version3.6 Page 3
Copyright (c) 1982 - 1998 by Assessment Systems

Conventional Item andest Analysis Program

Item analysis for data from file :\Program Files\ITEMAN\ali.dat
Date :May 16 2008 Time : 0:48 AM

Seq No.	Scale -	Prop. Item	Disc. Correct	Point Index	Alt.	Prop Total	Endorsing Low	High	Point	Key
		Item Statistics			Alternative					
8	1-8	.60	1.00	.77	1	.20	.00	.00	-.25	
					2	.10	.50	.00	-.41	
					3	.10	.50	.00	-.51	
					4	.60	.00	1.00	.77	*
					Other	.00	.00	.00		
9	1-9	.70	1.00	.81	1	.10	.50	.00	-.41	
					2	.70	.00	1.00	.81	*
					3	.10	.50	.00	-.51	
					4	.10	.00	.00	-.31	
					Other	.00	.00	.00		
10	1-10	.80	1.00	.69	1	.00	.00	.00		
					2	.10	.50	.00	-.41	
					3	.10	.50	.00	-.51	
					4	.80	.00	1.00	.69	*
					Other	.00	.00	.00		
11	2-1	.60	1.00	.40	1	.60	.00	1.00	.40	*
					2	.40	1.00	.00	-.40	
					Other	.00	.00	.00		
12	2-2	.70	1.00	.58	1	.70	.00	1.00	.58	*
					2	.30	1.00	.00	-.58	
					Other	.00	.00	.00		
13	2-3	.50	.17	.44	1	.50	.50	.33	-.44	
					2	.50	.50	.67	.44	*
					Other	.00	.00	.00		
14	2-4	.50	1.00	.63	1	.50	.00	1.00	.63	*
					2	.50	1.00	.00	-.63	
					Other	.00	.00	.00		
15	2-5	.60	1.00	.79	1	.40	1.00	.00	-.79	
					2	.60	.00	1.00	.79	*
					Other	.00	.00	.00		
16	2-6	.60	.17	.20	1	.40	.50	.33	-.20	
					2	.60	.50	.67	.20	*
					Other	.00	.00	.00		

الصفحة الرابعة:

```
 ITEMAN(tm ) for32-bit Windows   Version3.6        Page 4
        Copyright(c )1982 - 1998             by Assessment Systems

                        Conventional Item and Test Analysis

   Item analysis for datam file C:\Program Files\ITEMAN\ali.dat
Date :May 16 ,2008                                         Time : 0:48 AM



                          Item Statistics                Alternative
                  -----------------------  -------------------------------- ----
Seq .Scale   Prop .  Disc .      Point              Endorsing   Point
No .   -    Item  Correct  Index     .  Alt .       Total  Low   High   .Key
---- -----  -------  ------  ------  ----- -----  ----  ----  ------ ---

 17   2-7     .60     -.33    .10       1     .40    .00   .33   -.10
                                        2     .60   1.00   .67    .10    *
                                      Other   .00    .00   .00

 18   2-8     .70      .00    .16       1     .70   1.00  1.00    .16    *
                                        2     .30    .00   .00   -.16
                                      Other   .00    .00   .00

 19   2-9     .40      .67    .50       1     .40    .00   .67    .50    *
                                        2     .60   1.00   .33   -.50
                                      Other   .00    .00   .00

 20   2-10    .30      .33    .48       1     .70   1.00   .67   -.48
                                        2     .30    .00   .33    .48    *
                                      Other   .00    .00   .00
```

الصفحة الخامسة:

```
 ITEMAN(tm ) for32-bit Windows   Version3.6    Page 5
        Copyright(c )1982 - 1998             by Assessment Systems

                     Conventional Item and Test Analysis

Item analysis for data from file :Program Files\ITEMAN\ali.dat
Date :May 16  2008                                      Time : 0:48 AM


 There were10     examinees in the data .

Scale Statistics
----------------

  Scale :            1        2
                  ------- -------
 N of Items         10       10
 N of Examinees     10       10
Mean              6.200    5.500
Variance         11.560    4.250
Std .Dev .        3.400    2.062
Skew             -0.172    0.685
Kurtosis         -1.601   -0.184
Minimum           1.000    3.000
Maximum          10.000   10.000
Median            5.000    5.000
Alpha             0.894    0.502
SEM               1.108    1.455
Mean P            0.620    0.550
Mean Item-Tot .   0.716    0.426
Mean Biserial     0.902    0.545
Max Scor e(Low )      2        3
 N(Low Group )        2        2
Min Score (High )    10        7
 N(High Group )       3        3

                 Scale
---------------------

       1       2

  1  1.000 -0.371
  2 -0.371  1.000
```

الصفحة السادسة:

ITEMAN (tm) for32-bit Windows . Version3.6 Page 6
Copyright (c) 1982 - 1998 by Assessment Systems Corporation

Conventional Item and Test Analysis Program

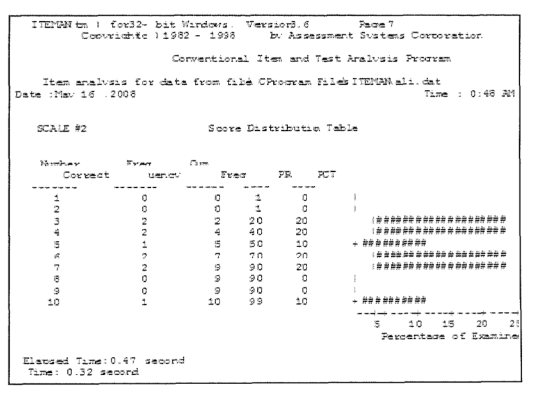
Item analysis for data from file C :\Program Files\ITEMAN\ali.dat
Date :May 16 .2008 Time : 0:48 AM

SCALE #1 Score Distribution Table

Number Correct	Freq - uency	Cum Freq	PR	PCT	
1	1	1	10	10	|##########
2	1	2	20	10	|##########
3	1	3	30	10	|##########
4	1	4	40	10	|##########
5	1	5	50	10	+##########
6	0	5	50	0	|
7	0	5	50	0	|
8	1	6	60	10	|##########
9	1	7	70	10	|##########
10	3	10	99	30	+##############################>

5 10 15 20 25
Percentage of Examinees

الصفحة السابعة:



ITEMAN (tm) for32- bit Windows . Version3.6 Page 7
Copyright (c) 1982 - 1998 by Assessment Systems Corporation

Conventional Item and Test Analysis Program

Item analysis for data from file CProgram FilesITEMANali.dat
Date :May 16 .2008 Time : 0:48 AM

SCALE #2 Score Distribution Table

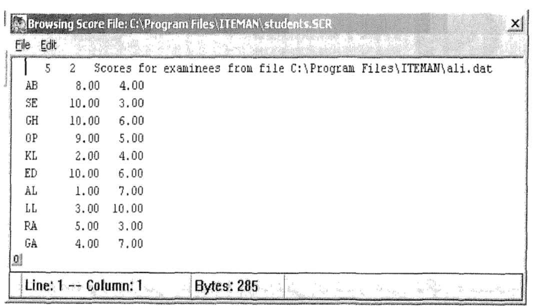
Number Correct	Freq uency	Cum Freq	PR	PCT	
1	0	0	1	0	|
2	0	0	1	0	|
3	2	2	20	20	|####################
4	2	4	40	20	|####################
5	1	5	50	10	+##########
6	2	7	70	20	|####################
7	2	9	90	20	|####################
8	0	9	90	0	|
9	0	9	90	0	|
10	1	10	99	10	+##########

5 10 15 20 25
Percentage of Examinees

Elapsed Time: 0.47 second
Time: 0.32 second

نتائج الطلبة:

أما نتيجة كل طالب في أسئلة الاختيار من متعدد وفي أسئلة الصواب والخطأ، فيوضحها الجدول الآتي:



AB	8.00	4.00
SE	10.00	3.00
GH	10.00	6.00
OP	9.00	5.00
KL	2.00	4.00
ED	10.00	6.00
AL	1.00	7.00
LL	3.00	10.00
RA	5.00	3.00
GA	4.00	7.00

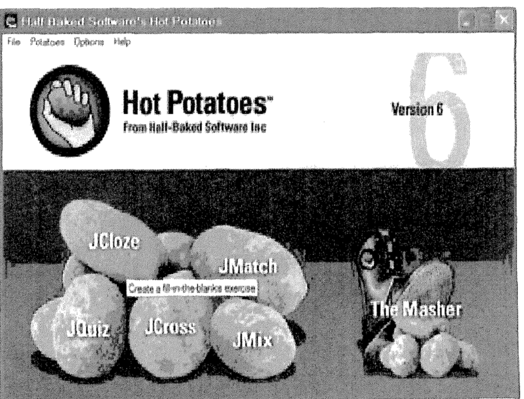
ثالثاً - برنامج Hot Potatoes





ما هو برنامج Hot Potatoes ؟ :

وهو أحد البرامج المعدة لغرض تصميم اختبار او استبيان داخل صفحة ويب بشكل سريع. ويمكن تعريفه بأنه مجموعة تطبيقية من البرامج المنشورة التي تم تصميمها بالتعاون ما بين جامعة فكتوريا وبرامج Half-Baked Software ويستخدم المدرسون هذا البرنامج لإعداد وتصميم المواد التعليمية وبخاصة الاختبارات والتمارين. كل هذه المواد يمكن أن تبرز على شكل صفحات إنترنت، وهذه الصفحات بالإمكان تحميلها بكل بساطة من خلال هذا البرنامج.



فكرة عامة عن برنامج Hot Potatoes

تتضمن حقيبة البرنامج التطبيقات الآتية:

الوظيفة- الاستخدام	البرنامج
1- الاختيار من متعدد.	
2- أسئلة المزاوجة.	
3- التكملة وملء الفراغات	
4- الكلمات المتقاطعة.	
5- إعادة ترتيب.	

الرسوم

هذا البرنامج مجاني للذين يعملون مع المؤسّسات التعليمية غير الربحيّة التي تجعل صفحاتهم متاحة على الويب المموّلة إعلانيا. ويتعين ان يدفع المستخدمون الآخرون رخصة استخدام البرنامج، ويمكن معرفة التكاليف من خلال زيارة الرابط التالي: http://web.uvic.ca/hrd/hotpot/index.htm

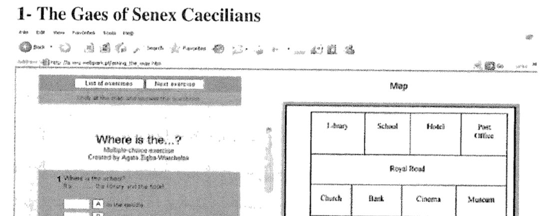
التسجيل

يشترط القيام بعملية التسجيل، وذلك حتى تتمكن من الاستخدام الأمثل للبرنامج. ويمكن القيام بذلك من خلال الرابط التالي:

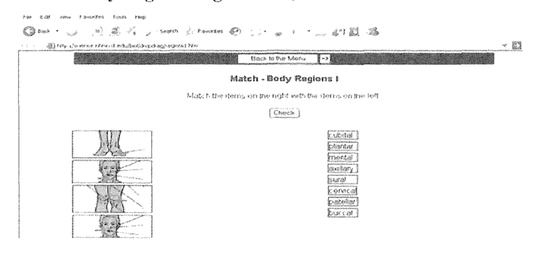
http://hotpot.uvic. ca/reg/register.htm

أمثلة على مواقع في الانترنت تم تصميمها باستخدام البرنامج :

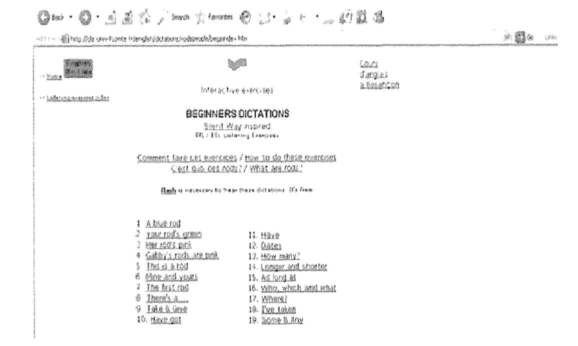
1- The Gaes of Senex Caecilians



2-Match- Body Regionsl(right to left)



3- Beginners dictations menu



واجهة البرنامج

ان برنامج Hot Potatoes مخصص لإعداد اختبارات وتمارين تفاعلية بلغة (html) وتحتوي واجهة البرنامج على ست برمجيات تمكن من كتابة أسئلة ذات خيارات متعددة، واسئلة تتطلب أجوبة قصيرة، وتمارين ذات فراغات،

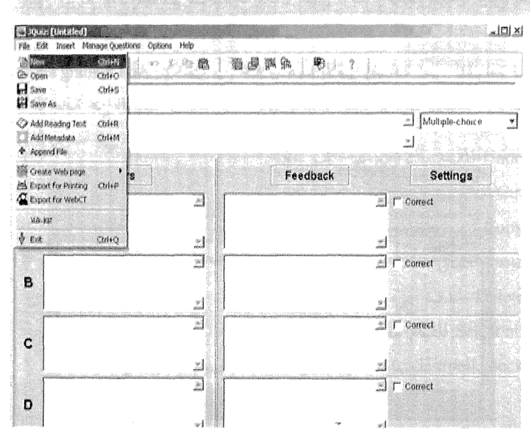
وتمارين الربط (المزاوجة)، وتمارين الكلمات المبعثرة، والكلمات المتقاطعة. ويبين الشكل التالي واجهة المستخدم الأساسية التي ستراها عند العمل في برنامج Hot Potatoes .

تحتوي كافة تطبيقات البرنامج على شريط مهام متشابه، وهو:

1- File (ملف).
2- Edit (تحرير).
3- Insert (إدراج).
4- Manage Items (إدارة الفقرات).
5- Option (الخيارات).
6- Help (المساعدة).

وفيما يلي شرح عن آلية الاستخدام:

(1) File



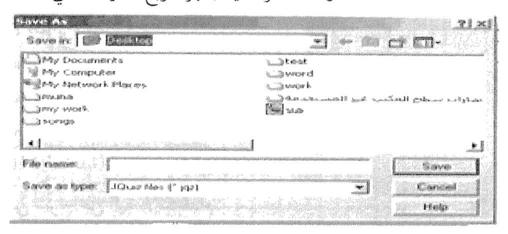
عند الضغط على File ستظهر مجموعة من الخيارات، وهي:

1- New لفتح إطار أو ملف جديد.

2- Open لفتح ملف معين موجود مسبقا.

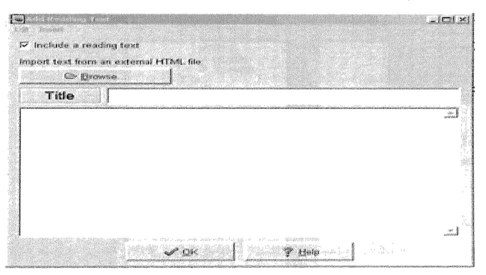
3- Save للقيام بحفظ التغييرات في الملف.

4- Save as لحفظ ملف، وعند النقر عليه يظهر مربع الحوار الآتي:

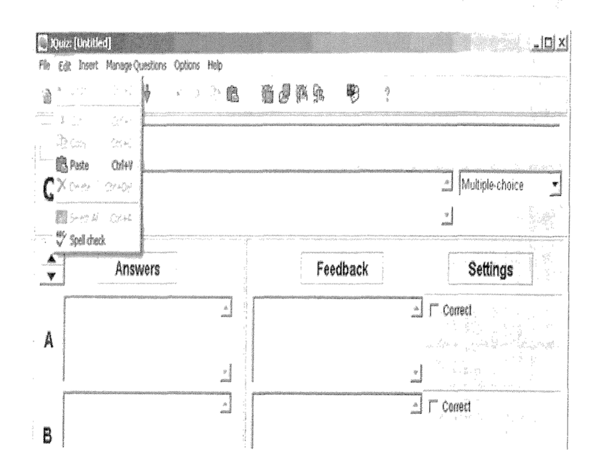


في مربع Save in قم بتحديد المكان الذي تريد الحفظ فيه، وفي file name قم بتحديد اسم الملف.

5- Add reading text يمكنك من خلال هذا الأمر إضافة نص للقراءة قبل الإجابة عن الأسئلة والتمارين، كما يمكنك أيضا توقيت القطعة (النص) المعروضة في فترة زمنية معينة، كما يبين الشكل التالي:



(2) Edit



تحتوي هذه القائمة على مجموعة من الأوامر، وهي:

1- Undo للقيام بالتراجع عن تغيير معين.

2- Cut لقطع جزء معين من نص.

3- Copy لنسخ قطعة معينة.

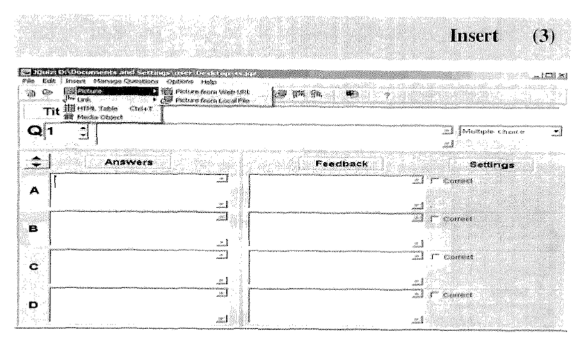
4- Paste للصق قطعة معينة.

5- Delete لحذف جزء معين.

6- Select all للقيام بتحديد كلي للنص.

7- Spell check للمساعدة في عملية القيام بالتصحيح الإملائي.

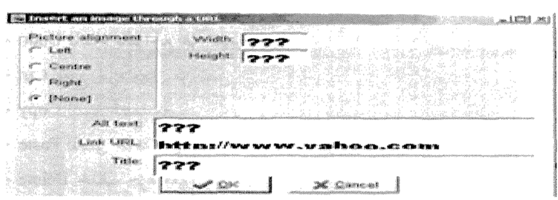
Insert (3)



تحتوي هذه القائمة على مجموعة من الاوامر وهي:

1- Picture يستخدم هذا الأمر عندما نريد إدراج صورة معينة، وعند الضغط على هذا الأمر سيظهر أمران:

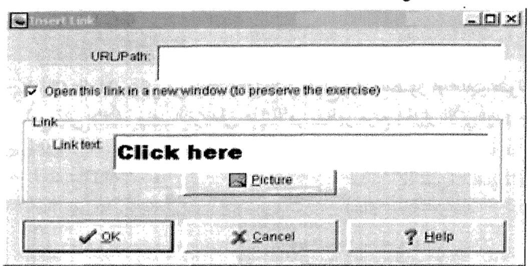
(أ) إدخال صورة من موقع، أو رابط معين Picture from web URL، وبعد النقر عليه سيظهر المربع المبين في الشكل أدناه، وبالتالي يمكنك كتابة اسم الرابط في link URL



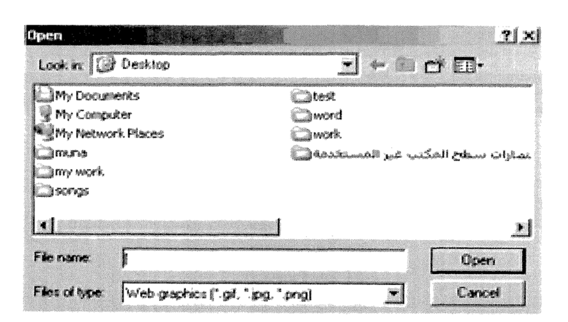
(ب) إدخال صورة من ملف معين Picture from file، وبعد النقر على Picture from file سيظهر المربع المبين في الشكل أدناه، وبالتالي يمكنك اختيار المكان الذي ستقوم بإدراج الصورة منه.

2 - link، يستخدم هذا الإمر عندما نريد إدراج رابطة معينة، وعند الضغط على هذا الأمر سيظهر أمران:

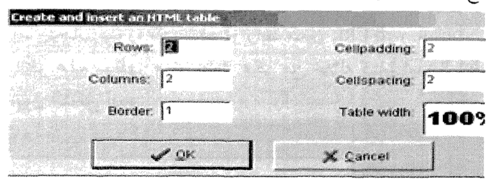
(أ) إما إدخال صورة من موقع أو رابط معين link to web URL، وبعد النقر عليه سيظهر مربع حوار يمكن من خلاله كتابة اسم الرابط كما يبين الشكل أدناه:



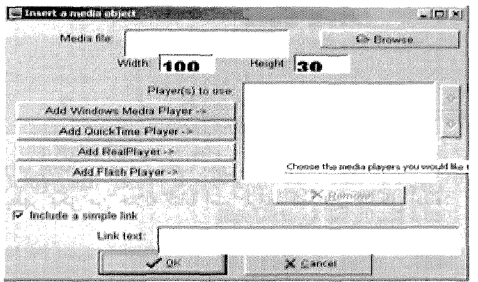
(ب) وإما إدخال رابط من ملف معين، وبعد النقر على link to local file سيظهر يمكن من خلاله كتابة اسم الرابط كما يبين الشكل أدناه:



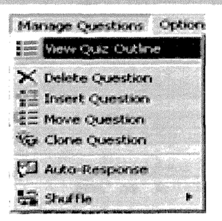
3- HTML table، من خلال هذا الأمر يمكنك إدراج جدول، وعند النقر عليه سيظهر مربع حوار يتضمن خصائص الجدول (عدد الصفوف والأعمدة) المطلوب إضافته.



4- Media Object، يمكن من خلال هذا الأمر إدراج ملف مصور او ملف صوت (موسيقى مثلاً)، وبعد النقر على هذا الأمر سيظهر مربع الحوار الآتي:



(4) Manage Items



تحتوي هذه القائمة على مجموعة من الأوامر وهي كالآتي:

1- View Quiz outline إظهار أساسيات أو مخطط الاختبار.

2- Delete Question حذف سؤال.

3- Insert Question: إدراج سؤال.

4- Move Question: نقل أو تحريك سؤال.

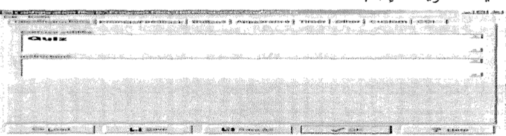
5- Clone Question: نسخ سؤال.

(5) Option

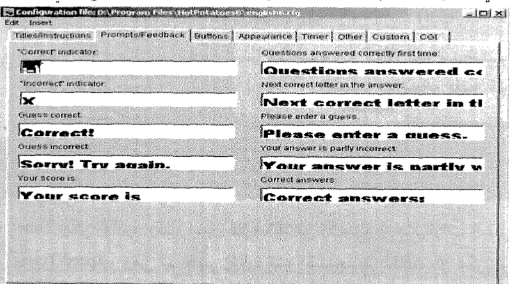
تحتوي هذه القائمة على مجموعة من الأوامر، ومن أهمها:

1- configure output، عند الضغط على هذا الامر سيظهر الشكل الآتي:

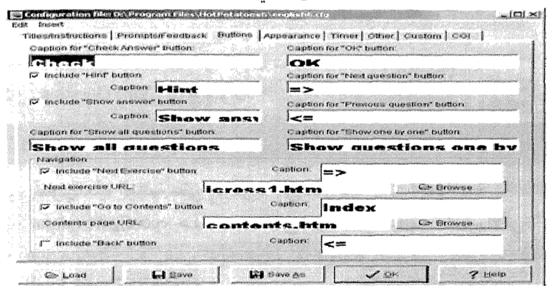
- Title/ Instruction يمكن كتابة عنوان الاختبار أو التمرين وبعض التعليمات المفيدة كطريقة الإجابة.



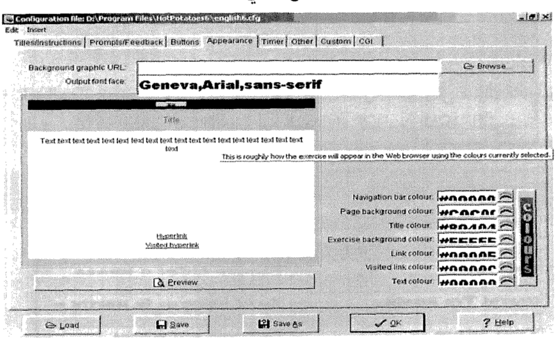
- Prompts/feedback يمكنك كتابة التغذية الراجعة والخاصة بكل سؤال (التعليق الذي سيقرؤه الطالب بعد الإجابة عن السؤال)، كما يبين الشكل الآتي:



- Buttons يمكن عمل ارتباط بين كل صفحة مصممة باستخدام البرنامج وصفحة أخرى، وكما يبين الشكل الآتي:

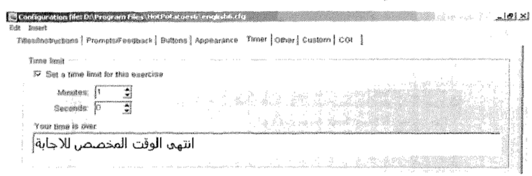


- Appearance يمكن القيام بتنسيق الصفحة باستخدام خطوط وألوان وخلفيات مختلفة، كما يبين الشكل الآتي:

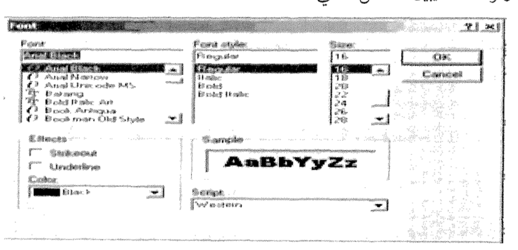


- Timer يمكن تحديد وقت معين للإجابة عن الأسئلة (اختبارات السرعة Speed Tests)، ومن ثم تظهر إشارة تنبه المستجيب بذلك، او يمكن تركه

إذا لم يكن الوقت عاملاً مهماً في الجواب (اختبارات القوة Power Tests)، كما يبين الشكل الآتي:



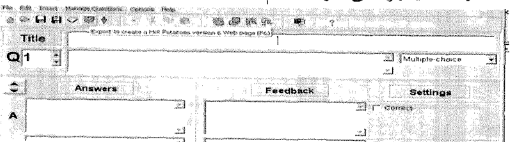
2- Font : عند النقر على هذا الأمر ستظهر قائمة تمكنك من تغيير نوع الخط ولونه كما يبين الشكل الآتي:



(6) Help

يستخدم هذا الأمر في مساعدة المستخدم في التعرف على البرنامج وطريقة عمله بخطوات مبسطة.

ملاحظة مهمة: يظهر على شريط المهام الإشارة التالية كما يبين الشكل أدناه:



عند الضغط على هذه الأيقونة يتم حفظ العمل، وتظهر ثلاثة خيارات، هي:

1- view the exercise in my browser إظهار التمرين على شاشة متصفحي

2- upload the file to the hotpotatoes. net website تحميل الملف إلى موقع البرنامج

3- nothing لا شي فقط يتم الحفظ

View or upload the exercise

The following file(s) have been created:
D:\Documents and Settings\user\Desktop\q.htm

What would you like to do?

View the exercise in my browser

Upload the file to the hotpotatoes.net Website

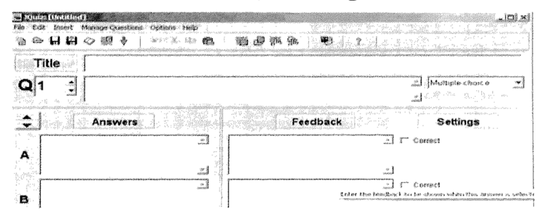
Nothing

Help

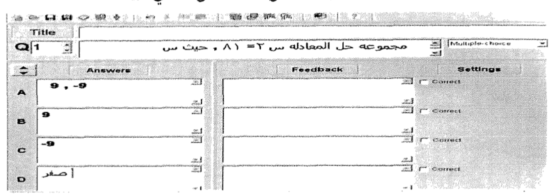
طريقة عمل برنامج Hot Potatoes

1 - طريقة عمل الإختيار من متعدد

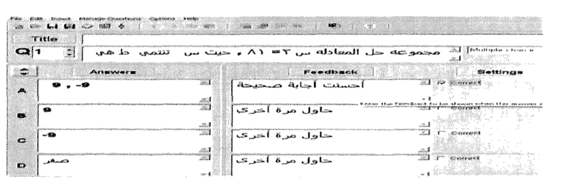
يتيح هذا الخيار إمكانية تصميم أسئلة من نوع الاختيار من متعدد، وبعد النقر على الأيقونة يظهر مربع الحوار الآتي:



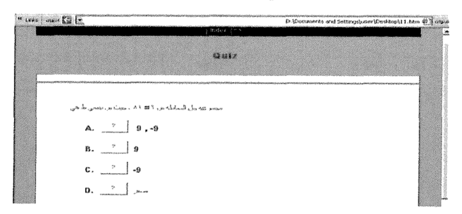
في مربع الحوار الخاص بالسؤال الأول (Q1) قم بكتابة جذر السؤال، وفي قائمة الإجابات (Answers) اكتب البدائل، والشكل التالي يبين ذلك:



يمكن الانتقال لكتابة سؤال آخر وذلك بالضغط على الأسهم الموجودة عند Q1، ويمكن أيضا تزويد السؤال بتغذية راجعة أو بمساعدة معينة وذلك في مربع حوار feedback كما يتوجب الإشارة إلى الإجابة الصحيحة في المربع الخاص في Setting كما يبين الشكل الآتي:



ومن ثم يمكن حفظ الملف وإظهاره كصفحة ويب من خلال الخطوات السابقة الذكر كما يبين الشكل الآتي:



2 – طريقة عمل اسئلة المزاوجة

يستخدم لإنشاء تمارين المزاوجة، وبعد النقر عليه يظهر مربع الحوار الآتي:

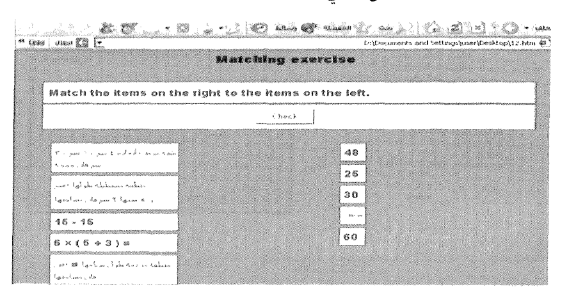
		Left (ordered) items	Right (jumbled) items
Title			
	1		
	2		Enter the text that matches the item
	3		
	4		
	5		
		Default:	???

في مربع الحوار الخاص بـ: Title أكتب عنوان الموضوع أو السؤال، وفي قائمة left (ordered) item أكتب الأسئلة، وفي العمود المجاور Right (jumbled) item لها أكتب الخيارات البديلة للإجابة كما يبين الشكل الآتي:

		Left (ordered) items	Right (jumbled) items	Fix
Title				
	1	شبه مكعب أبعاده ٤ سم ، ٥ سم ، ٣سم فان حجمه	48	
	2	منطقة مستطيلة طولها ٥سم وعرضها ٦ سم فان مساحتها	صفر	
	3	16 - 16	60	
	4	6 × (6 + 3) =	25	
	5	منطقة مربعة طول ضلعها = ٥س فان مساحتها	30	
		Default:	20	

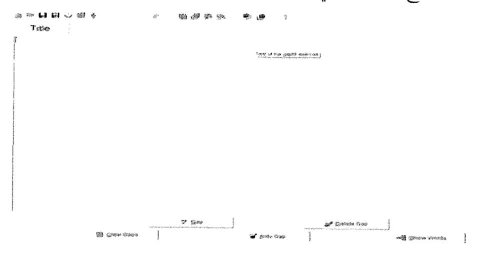
Enter the text that matches the item on the left

بعد ذلك، يمكن حفظ الملف وعرضه كصفحة ويب طبقاً للخطوات السابقة، وكما يبين الشكل الآتي:



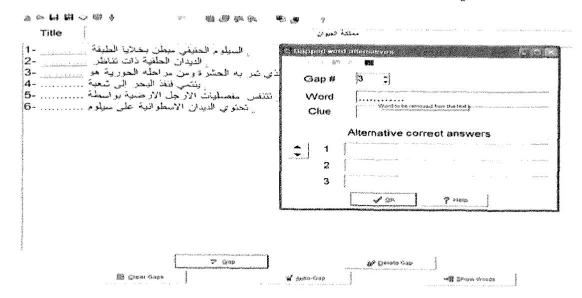
3 - طريقة عمل أسئلة التكملة وملء الفراغات

يمكن من خلال هذا الخيار إعداد أسئلة التكملة وملء الفراغات؛ حيث نقوم بكتابة الأسئلة او قطعة معينة، ونحدد الفراغات المراد إكمالها. وبعد النقر عليه يظهر مربع الحوار الآتي:



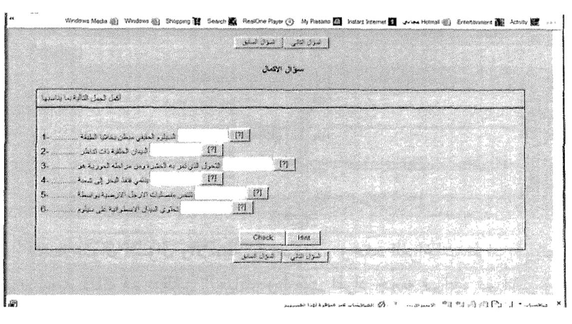
أكتب عنوان السؤال والأسئلة الخاصة به، وحدد الفراغات المراد إكمالها فيه، بعد ذلك قم بتظليل النقاط أو المكان المراد جعله فراغ، ثم أنقر زر

Gap ، سيتم فتح مربع حوار صغير أسمه Gapped word alternatives، وكما يبين الشكل الآتي:



بعد الانتهاء من عمل الفراغات، يمكن التحكم في تغيير كتابة التغذية الراجعة، كما يمكن التحكم بالوان الكتابة والخلفية لصفحة الويب التي سيظهر عليها السؤال فيما بعد.

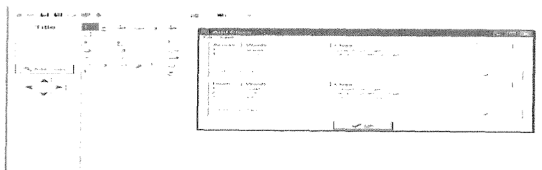
سيكون الشكل النهائي لهذا النوع من الأسئلة بعد نشره في الانترنت على الشكل الآتي:



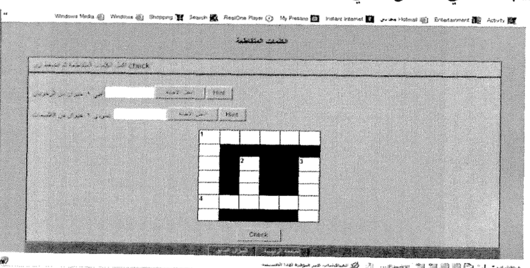
4 - طريقة عمل الكلمات المتقاطعة

يمكن تصميم أسئلة على شكل كلمات متقاطعة، ويتم كتابة الكلمات عموديا وأفقيا كما هو معد لذلك، أو يمكن إدخال الكلمات وتحديدها في الجزء العمودي أو الأفقي ويقوم البرنامج بمزجها. وبعد النقر على الايقونة يتم فتح مربع حوار، اكتب فيه الكلمات بشكل عمودي وأفقي، مع مفاتيح الإجابة في الاماكن المخصصة لذلك.

وبعد النقر على Add Clue يظهر مربع حوار يمكن من خلاله كتابة الكلمات في المربعات التي تظهر، ومفاتيح الإجابات، ويقوم البرنامج بمزجها. كما يمكننا تغيير لغة التعليمات وكتابة التعليمات، والتغذية الراجعة، والتحكم في ألوان الخلفيات، والكلمات فى صفحة الويب التى سنحصل عليها.



وأخيرا يمكن الحصول على سؤال الكلمات المتقاطعة على شكل صفحة ويب كما في الشكل الآتي:



5 - طريقة عمل إعادة الترتيب

يمكن تصميم أسئلة تحتوي على أجزاء مبعثرة (أجزاء من صور، اسماء، تواريخ، .... الخ)، ونطلب من الطالب ترتيبها بشكل صحيح. وبعد النقر على الأيقونة يظهر صندوق حوار، أكتب عنوان السؤال، والسؤال في الأماكن المخصصة، وكما موضح في الشكل الآتي:

Title ترتيب الشعب

Main sentence

Alternate sentences

1

2

3

Options for alternate sentences

Allow sentences which do not use all words and punctuation in main sentence

وبعد الانتهاء من ذلك، يكون الشكل النهائي للسؤال كما يلي:

Check Restart Hint

كلمة أخيرة:

إن برنامج Hot Potatoes شأنه - شأن برامج الحاسوب الأخرى - تعلمه سهل وسريع، ولكن التمكن منه واستخدامه بكفاءة يتوقف على كمية الممارسة والتدريب المستمر، كما أن التعامل معه على أنه لعبة مسلية تجعله ممتعاً، وتؤدي بالمستخدم إلى تعلم الكثير من البرنامج.

ننصح المعلم، والطالب، والباحث، تطبيق الخطوات الموضحة سابقاً في اختبار تحصيلي بمجال التخصص ونشره على شبكة الانترنت، والطلب من المستخدمين تقديم الملاحظات والتعليقات، سيؤدي إلى تطوّر الاختبار والتمكن من البرنامج.

مراجع الفصل العاشر

1- أبو سريع، رضا عبد الله (2004). تحليل البيانات باستخدام SPSS. عمان: دار الفكر.

2- الزعبي، محمد بلال؛ والطلافح، عباس (2003). النظام الإحصائي SPSS: فهم وتحليل البيانات الإحصائية. عمان: دار وائل للطباعة والنشر.

3- الضحيان، سعود بن ضحيان؛ وحسن، عزت عبد الحميد (2002). معالجة البيانات باستخدام برنامج SPSS 10.0. الرياض: مطابع التقنية.

4- العتوم، شفيق (2006). طرق الإحصاء: تطبيقات اقتصادية وإدارية باستخدام SPSS. عمان: دار المناهج.

5- عيد، غادة خالد (2006). القياس والتقويم التربوي مع تطبيقات برنامج SPSS. الكويت: مكتبة الفلاح للنشر والتوزيع.

6- المنيزل، عبد الله فلاح (2000). الإحصاء الاستدلالي وتطبيقاته في الحاسوب باستخدام الرزم الإحصائية SPSS. عمان: دار وائل للطباعة والنشر.

7- النجار، عبد الله بن عمر (2003). استخدام حزمة البرامج الإحصائية SPSS في تحليل البيانات. الرياض: مؤسسة شبكة البيانات.

8- نجيب، حسين علي؛ والرفاعي، غالب عوض صالح (2006). تحليل ونمذجة البيانات باستخدام الحاسوب، تطبيق شامل للحزمة SPSS. عمان: الأهلية للنشر والتوزيع.

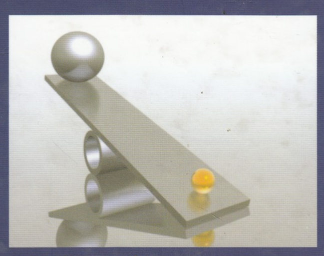
9- Assessment Systems Corporation (1996). **User's Manual for the ITEMAN Conventional Item Analysis Program**. Minnersota, USA.

10- Brace, N., Kemp, R, & Snelgar, R. (2000). **SPSS for Psychologists**. London: Macmillan Press LTD.

11- George, D., & Mallery, P. (2003). SPSS for Windows, Step by Step: A Simple **Guide and Reference 11.0 Update (4th ed.).** Boston: Allyn and Bacon.

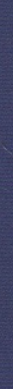
12- Kinear, P.R. (1994). **SPSS for Windows Made Simple**. London: Lawrence Erlbaum Associates.

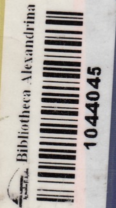
13- SPSS (2004). **SPSS 13.0 Brief Guide**. Ghicago: Polar Engineering and Consulting.



مفاهيم وتطبيقات

في التقويم والقياس التربوي







دولة الكويت

حولي - شارع بيروت - عمارة الأطباء

هاتف: 1985 264 فاكس: 7784 264 965+

ص.ب: 4848 الصفاة - الرمز البريدي 13049

دولة الإمارات العربية المتحدة

العين: - ص.ب 16431 هاتف: 7662189 فاكس: 7657901 3 971+

دبي: - ص.ب: 20438 هاتف: 2630618 فاكس: 2630628 4 971+

جمهورية مصر العربية

37 شارع النصر - امتداد رمسيس 2

مقابل وزارة المالية - مدينة نصر - القاهرة

هاتف: 22628143 فاكس: 22636587 02 002

www.alfalahbookshop.com

المطابع المركزية
عمان - الأردن